

Infinitesimals without Logic

Paolo Giordano¹

¹*Università della Svizzera italiana**

We introduce the ring of Fermat reals, an extension of the real field containing nilpotent infinitesimals. The construction takes inspiration from Smooth Infinitesimal Analysis (SIA), but provides a powerful theory of actual infinitesimals without any need of a background in mathematical logic. In particular, on the contrary with respect to SIA, which admits models only in intuitionistic logic, the theory of Fermat reals is consistent with classical logic. We face the problem to decide if the product of powers of nilpotent infinitesimals is zero or not, the identity principle for polynomials, the definition and properties of the total order relation. The construction is highly constructive, and every Fermat real admits a clear and order preserving geometrical representation. Using nilpotent infinitesimals, every smooth functions becomes a polynomial because in Taylor's formulas the rest is now zero. Finally, we present several applications to informal classical calculations used in Physics: now all these calculations become rigorous and, at the same time, formally equal to the informal ones. In particular, an interesting rigorous deduction of the wave equation is given, that clarifies how to formalize the approximations tied with Hook's law using this language of nilpotent infinitesimals.

Contents

I. Introduction and general problem	3
II. Motivations for the name “Fermat reals”	5
III. Definition and algebraic properties of Fermat reals: The basic idea	6
IV. First properties of little-oh polynomials	9

*Electronic address: paolo.giordano@usi.ch

	2
Little-oh polynomials are nilpotent:	9
Closure of little-oh polynomials with respect to smooth functions:	10
V. Equality and decomposition of Fermat reals	11
VI. The ideals D_k	15
VII. Products of powers of nilpotent infinitesimals	17
VIII. Identity principle for polynomials and invertible Fermat reals	19
IX. The derivation formula	21
X. Nilpotent infinitesimals and order properties	23
XI. Order relation	25
XII. Absolute value, powers and logarithms	32
XIII. Geometrical representation of Fermat reals	33
XIV. Some elementary examples	37
A. The heat equation.	38
B. Electric dipole.	40
C. Newtonian limit in Relativity.	41
D. Linear differential equations.	41
E. Circle of curvature.	42
F. Commutation of differentiation and integration.	42
G. Schwarz's theorem.	43
H. Area of the circle and volumes of revolution.	44
I. Curvature.	45
J. Stretching of a spring (and center of pressure).	46
K. The wave equation.	47
XV. Conclusions	59
References	60

I. INTRODUCTION AND GENERAL PROBLEM

Frequently in work by physicists it is possible to find informal calculations like

$$\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1 + \frac{v^2}{2c^2} \quad \sqrt{1 - h_{44}(x)} = 1 - \frac{1}{2}h_{44}(x) \quad (1)$$

with explicit use of infinitesimals $v/c \ll 1$ or $h_{44}(x) \ll 1$ such that e.g. $h_{44}(x)^2 = 0$. For example Einstein [13] wrote the formula (using the equality sign and not the approximate equality sign \simeq)

$$f(x, t + \tau) = f(x, t) + \tau \cdot \frac{\partial f}{\partial t}(x, t) \quad (2)$$

justifying it with the words “*since τ is very small*”; the formulas (1) are a particular case of the general (2). Dirac [10] wrote an analogous equality studying the Newtonian approximation in general relativity.

Using this type of infinitesimals we can write an *equality*, in some infinitesimal neighborhood, between a smooth function and its tangent straight line, or, in other words, a Taylor’s formula without remainder.

There are obviously many possibilities to formalize this kind of intuitive reasonings, obtaining a more or less good dialectic between informal and formal thinking, and indeed there are several theories of actual infinitesimals (from now on, for simplicity, we will say “infinitesimals” instead of “actual infinitesimals” as opposed to “potential infinitesimals”). Starting from these theories we can see that we can distinguish between two type of definitions of infinitesimals: in the first one we have at least a ring R containing the real field \mathbb{R} and infinitesimals are elements $\varepsilon \in R$ such that $-r < \varepsilon < r$ for every positive standard real $r \in \mathbb{R}_{>0}$. The second type of infinitesimal is defined using some algebraic property of nilpotency, i.e. $\varepsilon^n = 0$ for some natural number $n \in \mathbb{N}$. For some ring R these definitions can coincide, but anyway they lead, of course, only to the trivial infinitesimal $\varepsilon = 0$ if $R = \mathbb{R}$.

However these definitions of infinitesimals correspond to theories which are completely different in nature and underlying ideas. Indeed these theories can be seen in a more interesting way to belong to two different classes. In the first one we can put theories that need a certain amount of non trivial results of mathematical logic, whereas in the second one we have attempts to define sufficiently strong theories of infinitesimals without the use of non trivial results of mathematical logic. In the first class we have Non-Standard Analysis

(NSA) and Synthetic Differential Geometry (SDG, also called Smooth Infinitesimal Analysis, see e.g. Bell [3], Kock [20], Lavendhomme [22], Moerdijk and Reyes [23]), in the second one we have, e.g., Weil functors (see Kriegl and Michor [21]), Levi-Civita fields (see Berz [7], Shamseddine [25]), surreal numbers (see Conway [9], Ehresmann [12]), geometries over rings containing infinitesimals (see Bertram [6]). More precisely we can say that to work in NSA and SDG one needs a formal control deeply stronger than the one used in “standard mathematics”. Indeed to use NSA one has to be able to formally write the sentences one needs to use the transfer theorem. Whereas SDG does not admit models in classical logic, but in intuitionistic logic only, and hence we have to be sure that in our proofs there is no use of the law of the excluded middle, or e.g. of the classical part of De Morgan’s law or of some form of the axiom of choice or of the implication of double negation toward affirmation and any other logical principle which is not valid in intuitionistic logic. Physicists, engineers, but also the greatest part of mathematicians are not used to have this strong formal control in their work, and it is for this reason that there are attempts to present both NSA and SDG reducing as much as possible the necessary formal control, even if at some level this is technically impossible (see e.g. Henson [19], and Benci and Di Nasso [4, 5] for NSA; Bell [3] and Lavendhomme [22] for SDG, where using an axiomatic approach the authors try to postpone the very difficult construction of an intuitionistic model of a whole set theory using Topos).

On the other hand NSA is essentially the only theory of infinitesimals with a discrete diffusion and a sufficiently great community of working mathematicians and published results in several areas of mathematics and its applications, see e.g. Albeverio et al. [1]. SDG is the only theory of infinitesimals with non trivial, new and published results in differential geometry concerning infinite dimensional spaces like the space of all the diffeomorphisms of a generic (e.g. non compact) smooth manifold. In NSA we have only few results concerning differential geometry. Other theories of infinitesimals have not, at least up to now, the same formal strength of NSA or SDG or the same potentiality to be applied in several different areas of mathematics.

Our main aim, of which the present work represents a first step, is to find a theory of infinitesimals within “standard mathematics” (in the precise sense explained above of a formal control more “standard” and not so strong as the one needed e.g. in NSA or SDG) with results comparable with those of SDG, without forcing the reader to learn a strong

formal control of the mathematics he/she is doing. Because it has to be considered inside “standard mathematics”, our theory of infinitesimals must be compatible with classical logic.

Concretely, the idea of the present work is to by-pass the impossibility theorem about the incompatibility of SDG with classical logic that forces SDG to find models within intuitionistic logic.

Another point of view about present theories of infinitesimals is that, in spite of the fact that frequently they are presented using opposed motivations, they lack the intuitive interpretation of what the powerful formalism permits to do. For some concrete example in this direction, see Giordano [16]. Another aim of the present work is to construct a new theory of infinitesimals preserving always a very good dialectic between formal properties and intuitive interpretation.

More technically we want to show that it is possible to extend the real field adding nilpotent infinitesimals, arriving at an enlarged real line $\bullet\mathbb{R}$, by means of a very simple construction completely inside “standard mathematics”. Indeed to define the extension $\bullet\mathbb{R} \supset \mathbb{R}$ we shall use elementary analysis only. To avoid misunderstandings it is important to clarify that the purpose of the present work is not to give an alternative foundation of differential and integral calculus (like NSA), but to obtain a theory of nilpotent infinitesimals as a first step for the foundation of a smooth (\mathcal{C}^∞) differential geometry. For some preliminary results in this direction, see Giordano [16].

II. MOTIVATIONS FOR THE NAME “FERMAT REALS”

It is well known that historically two possible reductionist constructions of the real field starting from the rationals have been made. The first one is Dedekind’s order completion using sections of rationals, the second one is Cauchy’s metric space completion. Of course there are no historical reason to attribute our extension $\bullet\mathbb{R} \supset \mathbb{R}$ of the real field, to be described below, to Fermat, but there are strong motivations to say that, probably, he would have liked the underlying spirit and some properties of our theory. For example:

1. a formalization of Fermat’s infinitesimal method to derive functions is provable in our theory. We recall that Fermat’s idea was, roughly speaking and not on the basis of an accurate historical analysis which goes beyond the scope of the present work (see e.g.

Edwards [11], Eves [14]), to suppose first $h \neq 0$, to construct the incremental ratio

$$\frac{f(x+h) - f(x)}{h}$$

and, after suitable simplifications (sometimes using infinitesimal properties), to take in the final result $h = 0$.

2. Fermat's method to find the maximum or minimum of a given function $f(x)$ at $x = a$ was to take e to be extremely small so that the value of $f(x+h)$ was approximately equal to that of $f(x)$. In modern, algebraic language, it can be said that $f(x+h) = f(x)$ only if $h^2 = 0$, that is if e is a first order infinitesimal. Fermat was aware that this is not a "true" equality but some kind of approximation (ibidem). We will follow a similar idea to define $\bullet\mathbb{R}$ introducing a suitable equivalence relation to represent this equality.
3. Fermat has been described by Bell [2] as "the king of amateurs" of mathematics, and hence we can suppose that in its mathematical work the informal/intuitive part was stronger with respect to the formal one. For this reason we can think that he would have liked our idea to obtain a theory of infinitesimals preserving always the intuitive meaning and without forcing the working mathematician to be too much formal.

For these reason we chose the name "Fermat reals" for our ring $\bullet\mathbb{R}$ (note: without the possessive case, to underline that we are not attributing our construction of $\bullet\mathbb{R}$ to Fermat).

III. DEFINITION AND ALGEBRAIC PROPERTIES OF FERMAT REALS: THE BASIC IDEA

We start from the idea that a smooth (C^∞) function $f : \bullet\mathbb{R} \longrightarrow \bullet\mathbb{R}$ is actually *equal* to its tangent straight line in the first order neighborhood e.g. of the point $x = 0$, that is

$$\forall h \in D : f(h) = f(0) + h \cdot f'(0) \tag{3}$$

where D is the subset of $\bullet\mathbb{R}$ which defines the above-mentioned neighborhood of $x = 0$. The equality (3) can be seen as a first-order Taylor's formula without remainder because intuitively we think that $h^2 = 0$ for any $h \in D$ (indeed the property $h^2 = 0$ defines the

first order neighborhood of $x = 0$ in $\bullet\mathbb{R}$). These almost trivial considerations lead us to understand many things: $\bullet\mathbb{R}$ must necessarily be a ring and not a field because in a field the equation $h^2 = 0$ implies $h = 0$; moreover we will surely have some limitation in the extension of some function from \mathbb{R} to $\bullet\mathbb{R}$, e.g. the square root, because using this function with the usual properties, once again the equation $h^2 = 0$ implies $|h| = 0$. On the other hand, we are also led to ask whether (3) uniquely determines the derivative $f'(0)$: because, even if it is true that we cannot simplify by h , we know that the polynomial coefficients of a Taylor's formula are unique in classical analysis. In fact we will prove that

$$\exists! m \in \mathbb{R} \forall h \in D : f(h) = f(0) + h \cdot m \quad (4)$$

that is the slope of the tangent is uniquely determined in case it is an ordinary real number. We will call formulas like (4) *derivation formulas*.

If we try to construct a model for (4) a natural idea is to think our new numbers in $\bullet\mathbb{R}$ as equivalence classes $[h]$ of usual functions $h : \mathbb{R} \longrightarrow \mathbb{R}$. In this way we may hope both to include the real field using classes generated by constant functions, and that the class generated by $h(t) = t$ could be a first order infinitesimal number. To understand how to define this equivalence relation we have to think at (3) in the following sense:

$$f(h(t)) \sim f(0) + h(t) \cdot f'(0), \quad (5)$$

where the idea is that we are going to define \sim . If we think $h(t)$ “sufficiently similar to t ”, we can define \sim so that (5) is equivalent to

$$\lim_{t \rightarrow 0^+} \frac{f(h(t)) - f(0) - h(t) \cdot f'(0)}{t} = 0,$$

that is

$$x \sim y \quad : \Longleftrightarrow \quad \lim_{t \rightarrow 0^+} \frac{x_t - y_t}{t} = 0. \quad (6)$$

In this way (5) is very near to the definition of differentiability for f at 0.

It is important to note that, because of de L'Hôpital's theorem we have the isomorphism

$$\mathcal{C}^1(\mathbb{R}, \mathbb{R}) / \sim \simeq \mathbb{R}[x] / (x),$$

the left hand side is (isomorphic to) the usual tangent bundle of \mathbb{R} and thus we obtain nothing new. It is not easy to understand what set of functions we have to choose for x, y

in (6) so as to obtain a non trivial structure. The first idea is to take continuous functions at $t = 0$, instead of more regular ones like \mathcal{C}^1 -functions, so that e.g. $h_k(t) = |t|^{1/k}$ becomes a k -th order nilpotent infinitesimal ($h^{k+1} \sim 0$); indeed for almost all the results presented in this article, continuous functions at $t = 0$ work well. However, only in proving the non-trivial property

$$(\forall x \in \bullet\mathbb{R} : x \cdot f(x) = 0) \implies \forall x \in \bullet\mathbb{R} : f(x) = 0 \quad (7)$$

we can see that it does not suffice to take continuous functions at $t = 0$. To prove (7) the following functions turned out to be very useful:

Definition 1. If $x : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$, then we say that x is *nilpotent* iff $|x(t) - x(0)|^k = o(t)$ as $t \rightarrow 0^+$, for some $k \in \mathbb{N}$. \mathcal{N} will denote the set of all the nilpotent functions.

E.g. any Hoelder function $|x(t) - x(s)| \leq c \cdot |t - s|^\alpha$ (for some constant $\alpha > 0$) is nilpotent. The choice of nilpotent functions instead of more regular ones establish a great difference of our approach with respect to the classical definition of jets (see e.g. Bröcker [8], Golubitsky and Guillemin [17]), that (6) may recall.

Another problem necessarily connected with the basic idea (3) is that the use of nilpotent infinitesimals very frequently leads to consider terms like $h_1^{i_1} \cdot \dots \cdot h_n^{i_n}$. For this type of products the first problem is to know whether $h_1^{i_1} \cdot \dots \cdot h_n^{i_n} \neq 0$ and what is the order k of this new infinitesimal, that is for what k we have $(h_1^{i_1} \cdot \dots \cdot h_n^{i_n})^k \neq 0$ but $(h_1^{i_1} \cdot \dots \cdot h_n^{i_n})^{k+1} = 0$. We will have a good frame if we will be able to solve these problems starting from the order of each infinitesimal h_j and from the values of the powers $i_j \in \mathbb{N}$. On the other hand almost all the examples of nilpotent infinitesimals are of the form $h(t) = t^\alpha$, with $0 < \alpha < 1$, and their sums; these functions have also great properties in the treatment of products of powers. It is for these reasons that we shall focus our attention on the following family of functions $x : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$ in the definition (6) of \sim :

Definition 2. We say that x is a *little-oh polynomial*, and we write $x \in \mathbb{R}_o[t]$ iff

1. $x : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$

2. We can write

$$x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o(t) \quad \text{as} \quad t \rightarrow 0^+$$

for suitable

$$k \in \mathbb{N}$$

$$r, \alpha_1, \dots, \alpha_k \in \mathbb{R}$$

$$a_1, \dots, a_k \in \mathbb{R}_{\geq 0}$$

Hence a little-oh polynomial $x \in \mathbb{R}_o[t]$ is a polynomial function with real coefficients, in the real variable $t \geq 0$, with generic positive powers of t , and up to a little-oh function as $t \rightarrow 0^+$.

Remark 3. In the following, writing $x_t = y_t + o(t)$ as $t \rightarrow 0+$ we will always mean

$$\lim_{t \rightarrow 0^+} \frac{x_t - y_t}{t} = 0 \quad \text{and} \quad x_0 = y_0.$$

In other words, every little-oh function we will consider is continuous as $t \rightarrow 0^+$.

Example. Simple examples of little-oh polynomials are the following:

1. $x_t = 1 + t + t^{1/2} + t^{1/3} + o(t)$
2. $x_t = r \quad \forall t$. Note that in this example we can take $k = 0$, and hence α and a are the void sequence of reals, that is the function $\alpha = a : \emptyset \rightarrow \mathbb{R}$, if we think of an n -tuple x of reals as a function $x : \{1, \dots, n\} \rightarrow \mathbb{R}$.
3. $x_t = r + o(t)$

IV. FIRST PROPERTIES OF LITTLE-OH POLYNOMIALS

Little-oh polynomials are nilpotent:

First properties of little-oh polynomials are the following: if $x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o_1(t)$ as $t \rightarrow 0^+$ and $y_t = s + \sum_{j=1}^N \beta_j \cdot t^{b_j} + o_2(t)$, then $(x+y)_t = r+s + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + \sum_{j=1}^N \beta_j \cdot t^{b_j} + o_3(t)$ and $(x \cdot y)_t = rs + \sum_{i=1}^k s\alpha_i \cdot t^{a_i} + \sum_{j=1}^N r\beta_j \cdot t^{b_j} + \sum_{i=1}^k \sum_{j=1}^N \alpha_i \beta_j \cdot t^{a_i} t^{b_j} + o_4(t)$, hence the set of little-oh polynomials is closed with respect to pointwise sum and product. Moreover little-oh polynomials are nilpotent (see Definition 1) functions; to prove this we firstly prove that the set of nilpotent functions \mathcal{N} is a subalgebra of the algebra $\mathbb{R}^{\mathbb{R}}$ of real valued functions. Indeed, let x and y be two nilpotent functions such that $|x - x(0)|^k = o_1(t)$ and

$|y - y(0)|^N = o_2(t)$, then we can write $x \cdot y - x(0) \cdot y(0) = x \cdot [y - y(0)] + y(0) \cdot [x - x(0)]$, so that we can consider $|x \cdot [y - y(0)]|^k = |x|^k \cdot |y - y(0)|^k = |x|^k \cdot o_1(t)$ and $\frac{|x|^k \cdot o_1(t)}{t} \rightarrow 0$ as $t \rightarrow 0^+$ because $|x|^k \rightarrow |x(0)|^k$, hence $x \cdot [y - y(0)] \in \mathcal{N}$. Analogously $y(0) \cdot [x - x(0)] \in \mathcal{N}$ and hence the closure of \mathcal{N} with respect to the product follows from the closure with respect to the sum. The case of the sum follows from the following equalities (where we use $x_t := x(t)$, $u := x - x_0$, $v := y - y_0$, $|u_t|^k = o_1(t)$ and $|v_t|^N = o_2(t)$ and we have supposed $k \geq N$):

$$u^k = o_1(t), \quad v^k = o_2(t)$$

$$(u + v)^k = \sum_{i=0}^k \binom{k}{i} u^i \cdot v^{k-i}$$

$$\forall i = 0, \dots, k : \frac{u_t^i \cdot v_t^{k-i}}{t} = \frac{(u_t^k)^{\frac{i}{k}} \cdot (v_t^k)^{\frac{k-i}{k}}}{t^{\frac{i}{k}} \cdot t^{\frac{k-i}{k}}} = \left(\frac{u_t^k}{t}\right)^{\frac{i}{k}} \cdot \left(\frac{v_t^k}{t}\right)^{\frac{k-i}{k}}.$$

Now we can prove that $\mathbb{R}_o[t]$ is a subalgebra of \mathcal{N} . Indeed every constant $r \in \mathbb{R}$ and every power t^{a_i} are elements of \mathcal{N} and hence $r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} \in \mathcal{N}$, so it remains to prove that if $y \in \mathcal{N}$ and $w = o(t)$, then $y + w \in \mathcal{N}$, but this is a consequence of the fact that every little-oh function is trivially nilpotent, and hence it follows from the closure of \mathcal{N} with respect to the sum.

Closure of little-oh polynomials with respect to smooth functions:

Now we want to prove that little-oh polynomials are preserved by smooth functions, that is if $x \in \mathbb{R}_o[t]$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ is smooth, then $f \circ x \in \mathbb{R}_o[t]$. Let us fix some notations:

$$x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + w(t) \quad \text{with} \quad w(t) = o(t)$$

$$h(t) := x(t) - x(0) \quad \forall t \in \mathbb{R}_{\geq 0}$$

hence $x_t = x(0) + h_t = r + h_t$. The function $t \mapsto h(t) = \sum_{i=1}^k \alpha_i \cdot t^{a_i} + w(t)$ belongs to $\mathbb{R}_o[t] \subseteq \mathcal{N}$ so we can write $|h|^N = o(t)$ for some $N \in \mathbb{N}$ and as $t \rightarrow 0^+$. From Taylor's formula we have

$$f(x_t) = f(r + h_t) = f(r) + \sum_{i=1}^N \frac{f^{(i)}(r)}{i!} \cdot h_t^i + f(x_t) = f(r + h_t) \quad (8)$$

$$= f(r) + \sum_{i=1}^N \frac{f^{(i)}(r)}{i!} \cdot h_t^i + o(h_t^N) \quad (9)$$

But

$$\frac{|o(h_t^N)|}{|t|} = \frac{|o(h_t^N)|}{|h_t^N|} \cdot \frac{|h_t^N|}{|t|} \rightarrow 0$$

hence $o(h_t^N) = o(t) \in \mathbb{R}_o[t]$. From this, the formula (8), the fact that $h \in \mathbb{R}_o[t]$ and using the closure of little-oh polynomials with respect to ring operations, the conclusion $f \circ x \in \mathbb{R}_o[t]$ follows.

V. EQUALITY AND DECOMPOSITION OF FERMAT REALS

Definition 4. Let $x, y \in \mathbb{R}_o[t]$, then we say that $x \sim y$ or that $x = y$ in $\bullet\mathbb{R}$ iff $x(t) = y(t) + o(t)$ as $t \rightarrow 0^+$. Because it is easy to prove that \sim is an equivalence relation, we can define $\bullet\mathbb{R} := \mathbb{R}_o[t] / \sim$, i.e. $\bullet\mathbb{R}$ is the quotient set of $\mathbb{R}_o[t]$ with respect to the equivalence relation \sim .

The equivalence relation \sim is a congruence with respect to pointwise operations, hence $\bullet\mathbb{R}$ is a commutative ring. Where it will be useful to simplify notations we will write “ $x = y$ in $\bullet\mathbb{R}$ ” instead of $x \sim y$, and we will talk directly about the elements of $\mathbb{R}_o[t]$ instead of their equivalence classes; for example we can say that $x = y$ in $\bullet\mathbb{R}$ and $z = w$ in $\bullet\mathbb{R}$ imply $x + z = y + w$ in $\bullet\mathbb{R}$.

The immersion of \mathbb{R} in $\bullet\mathbb{R}$ is $r \mapsto \hat{r}$ defined by $\hat{r}(t) := r$, and in the sequel we will always identify $\hat{\mathbb{R}}$ with \mathbb{R} , which is hence a subring of $\bullet\mathbb{R}$. Conversely if $x \in \bullet\mathbb{R}$ then the map $^\circ(-) : x \in \bullet\mathbb{R} \mapsto ^\circ x = x(0) \in \mathbb{R}$, which evaluates each extended real in 0, is well defined. We shall call $^\circ(-)$ the *standard part map*. Let us also note that, as a vector space over the field \mathbb{R} we have $\dim_{\mathbb{R}} \bullet\mathbb{R} = \infty$, and this underlines even more the difference of our approach with respect to the classical definition of jets. Our idea is instead more near to NSA, where standard sets can be extended adding new infinitesimal points, and this is not the point of view of jet theory.

With the following theorem we will introduce the decomposition of a Fermat real $x \in \bullet\mathbb{R}$, that is a unique notation for its standard part and all its infinitesimal parts.

Theorem 5. *If $x \in \bullet\mathbb{R}$, then there exist one and only one sequence*

$$(k, r, \alpha_1, \dots, \alpha_k, a_1, \dots, a_k)$$

such that

$$k \in \mathbb{N}$$

$$r, \alpha_1, \dots, \alpha_k, a_1, \dots, a_k \in \mathbb{R}$$

and

1. $x = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i}$ in $\bullet\mathbb{R}$
2. $0 < a_1 < a_2 < \dots < a_k \leq 1$
3. $\alpha_i \neq 0 \quad \forall i = 1, \dots, k$

In this statement we have also to include the void case $k = 0$ and $\alpha = a : \emptyset \longrightarrow \mathbb{R}$. Obviously, as usual, we use the definition $\sum_{i=1}^0 b_i = 0$ for the sum of an empty set of numbers. As we shall see, this is the case where x is a standard real, i.e. $x \in \mathbb{R}$.

In the following we will use the notations $t^a := dt_{1/a} := [t \in \mathbb{R}_{\geq 0} \mapsto t^a \in \mathbb{R}]_{\sim} \in \bullet\mathbb{R}$ so that e.g. $dt_2 = t^{1/2}$ is a second order infinitesimal. In general, as we will see from the definition of order of a generic infinitesimal, dt_a is an infinitesimal of order a . In other words these two notations for the same object permit to emphasize the difference between an actual infinitesimal dt_a and a potential infinitesimal $t^{1/a}$: an actual infinitesimal of order $a \geq 1$ corresponds to a potential infinitesimal of order $\frac{1}{a} \leq 1$ (with respect to the classical notion of order of an infinitesimal function from calculus, see e.g. Prodi [24], Silov [26]).

Remark 6. Let us note that $dt_a \cdot dt_b = dt_{\frac{ab}{a+b}}$, moreover $dt_a^\alpha := (dt_a)^\alpha = dt_{\frac{a}{\alpha}}$ for every $\alpha \geq 1$ and finally $dt_a = 0$ for every $a < 1$. E.g. $dt_a^{[a]+1} = 0$ for every $a \in \mathbb{R}_{>0}$, where $[a] \in \mathbb{N}$ is the integer part of a , i.e. $[a] \leq a < [a] + 1$.

Existence proof:

Since $x \in \mathbb{R}_o[t]$, we can write $x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o(t)$ as $t \rightarrow 0^+$, where $r, \alpha_i \in \mathbb{R}$, $a_i \in \mathbb{R}_{\geq 0}$ and $k \in \mathbb{N}$. Hence $x = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i}$ in $\bullet\mathbb{R}$ and our purpose is to pass from this representation of x to another one that satisfies conditions 1, 2 and 3 of the statement. Since if $a_i > 1$ then $\alpha_i \cdot t^{a_i} = 0$ in $\bullet\mathbb{R}$, we can suppose that $a_i \leq 1$ for every $i = 1, \dots, k$. Moreover we can also suppose $a_i > 0$ for every i , because otherwise, if $a_i = 0$, we can replace $r \in \mathbb{R}$ by $r + \sum \{\alpha_i \mid a_i = 0, i = 1, \dots, k\}$.

Now we sum all the terms t^{a_i} having the same a_i , that is we can consider

$$\bar{\alpha}_i := \sum \{\alpha_j \mid a_j = a_i, j = 1, \dots, k\}$$

so that in $\bullet\mathbb{R}$ we have

$$x = r + \sum_{i \in I} \bar{\alpha}_i \cdot t^{a_i}$$

where $I \subseteq \{1, \dots, k\}$, $\{a_i \mid i \in I\} = \{a, \dots, a_k\}$ and $a_i \neq a_j$ for any $i, j \in I$ with $i \neq j$. Neglecting $\bar{\alpha}_i$ if $\bar{\alpha}_i = 0$ and renaming a_i , for $i \in I$, in such a way that $a_i < a_j$ if $i, j \in I$ with $i < j$, we obtain the existence result. Note that if $x = r \in \mathbb{R}$, in the final step of this proof we have $I = \emptyset$.

Uniqueness proof:

Let us suppose that in $\bullet\mathbb{R}$ we have

$$x = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} = s + \sum_{j=1}^N \beta_j \cdot t^{b_j} \quad (10)$$

where α_i, β_j, a_i and b_j verify the conditions of the statement. First of all ${}^\circ x = x(0) = r = s$ because $a_i, b_j > 0$. Hence $\alpha_1 t^{a_1} - \beta_1 t^{b_1} + \sum_i \alpha_i \cdot t^{a_i} - \sum_j \beta_j \cdot t^{b_j} = o(t)$. By reduction to the absurd, if we had $a_1 < b_1$, then collecting the term t^{a_1} we would have

$$\alpha_1 - \beta_1 t^{b_1-a_1} + \sum_i \alpha_i \cdot t^{a_i-a_1} - \sum_j \beta_j \cdot t^{b_j-a_1} = \frac{o(t)}{t} \cdot t^{1-a_1}. \quad (11)$$

In (11) we have that $\beta_1 t^{b_1-a_1} \rightarrow 0$ for $t \rightarrow 0^+$ because $a_1 < b_1$ by hypothesis; $\sum_i \alpha_i \cdot t^{a_i-a_1} \rightarrow 0$ because $a_1 < a_i$ for $i = 2, \dots, k$; $\sum_j \beta_j \cdot t^{b_j-a_1} \rightarrow 0$ because $a_1 < b_1 < b_j$ for $j = 2, \dots, N$, and finally t^{1-a_1} is limited because $a_1 \leq 1$. Hence for $t \rightarrow 0^+$ we obtain $\alpha_1 = 0$, which conflicts with condition 3 of the statement. We can argue in a corresponding way if we had $b_1 < a_1$. In this way we see that we must have $a_1 = b_1$. From this and from equation (11) we obtain

$$\alpha_1 - \beta_1 + \sum_i \alpha_i \cdot t^{a_i-a_1} - \sum_j \beta_j \cdot t^{b_j-a_1} = \frac{o(t)}{t} \cdot t^{1-a_1} \quad (12)$$

and hence for $t \rightarrow 0^+$ we obtain $\alpha_1 = \beta_1$. We can now restart from (12) to prove, in the same way, that $a_2 = b_2, \alpha_2 = \beta_2$, etc. At the end we must have $k = N$ because, otherwise, if we had e.g. $k < N$, at the end of the previous recursive process, we would have

$$\sum_{j=k+1}^N \beta_j \cdot t^{b_j} = o(t).$$

From this, collecting the terms containing $t^{b_{k+1}}$, we obtain

$$t^{b_{k+1}-1} \cdot [\beta_{k+1} + \beta_{k+2} \cdot t^{b_{k+2}-b_{k+1}} + \dots + \beta_N \cdot t^{\beta_N-\beta_{k+1}}] \rightarrow 0. \quad (13)$$

In this sum $\beta_{k+j} \cdot t^{b_{k+j}-b_{k+1}} \rightarrow 0$ as $t \rightarrow 0^+$, because $b_{k+1} < b_{k+j}$ for $j > 1$ and hence $\beta_{k+1} + \beta_{k+2} \cdot t^{b_{k+2}-b_{k+1}} + \dots + \beta_N \cdot t^{\beta_N-\beta_{k+1}} \rightarrow \beta_{k+1} \neq 0$, so from (13) we get $t^{b_{k+1}-1} \rightarrow 0$, that is $b_{k+1} > 1$, in contradiction with the uniqueness hypothesis $b_{k+1} \leq 1$.

Let us note explicitly that the uniqueness proof permits also to affirm that the decomposition is well defined in $\bullet\mathbb{R}$, i.e. that if $x = y$ in $\bullet\mathbb{R}$, then the decomposition of x and the decomposition of y are equal.

■

On the basis of this theorem we introduce two notations: the first one emphasizing the potential nature of an infinitesimal $x \in \bullet\mathbb{R}$, and the second one emphasizing its actual nature.

Definition 7. If $x \in \bullet\mathbb{R}$, we say that

$$x = r + \sum_{i=1}^k \alpha_i \cdot t^{\alpha_i} \text{ is the potential decomposition (of } x) \quad (14)$$

iff conditions 1., 2., and 3. of Theorem 5 are verified. Of course it is implicit that the symbol of equality in (14) has to be understood in $\bullet\mathbb{R}$.

For example $x = 1 + t^{1/3} + t^{1/2} + t$ is a decomposition because we have increasing powers of t . The only decomposition of a standard real $r \in \mathbb{R}$ is the void one, i.e. that with $k = 0$ and $\alpha = a : \emptyset \longrightarrow \mathbb{R}$; indeed to see that this is the case, it suffices to go along the existence proof again with this case $x = r \in \mathbb{R}$ (or to prove it directly, e.g. by contradiction).

Definition 8. Considering that $t^{\alpha_i} = dt_{1/a_i}$ we can also use the following notation, emphasizing more the fact that $x \in \bullet\mathbb{R}$ is an actual infinitesimal:

$$x = {}^\circ x + \sum_{i=1}^k {}^\circ x_i \cdot dt_{b_i} \quad (15)$$

where we have used the notation ${}^\circ x_i := \alpha_i$ and $b_i := 1/a_i$, so that the condition that uniquely identifies all b_i is $b_1 > b_2 > \dots > b_k \geq 1$. We call (15) the *actual decomposition* of x or simply the *decomposition* of x . We will also use the notation $d^i x := {}^\circ x_i \cdot dt_{b_i}$ (and simply $dx := d^1 x$) and we will call ${}^\circ x_i$ the *i-th standard part of x* and $d^i x$ the *i-th infinitesimal part of x* or the *i-th differential of x* . So let us note that we can also write

$$x = {}^\circ x + \sum_i d^i x$$

and in this notation all the addenda are uniquely determined (the number of them too). Finally, if $k \geq 1$ that is if $x \in \bullet\mathbb{R} \setminus \mathbb{R}$, we set $\omega(x) := b_1$ and $\omega_i(x) := b_i$. The real number $\omega(x) = b_1$ is the greatest order in the actual decomposition (15), corresponding to

the smallest in the potential decomposition (14), and is called the *order* of the Fermat real $x \in \bullet\mathbb{R}$. The number $\omega_i(x) = b_i$ is called the i -th order of x . If $x \in \mathbb{R}$ we set $\omega(x) := 0$ and $d^i x := 0$. Observe that in general $\omega(x) = \omega(dx)$, $d(dx) = dx$ and that, using the notations of the potential decomposition (7), we have $\omega(x) = 1/a_1$.

Example. If $x = 1 + t^{1/3} + t^{1/2} + t$, then ${}^\circ x = 1$, $dx = dt_3$ and hence x is a third order infinitesimal, i.e. $\omega(x) = 3$, $d^2 x = dt_2$ and $d^3 x = dt$; finally all the standard parts are ${}^\circ x_i = 1$.

VI. THE IDEALS D_k

In this section we will introduce the sets of nilpotent infinitesimals corresponding to a k -th order neighborhood of 0. Every smooth function restricted to this neighborhood becomes a polynomial of order k , obviously given by its k -th order Taylor's formula (without remainder). We start with a theorem characterizing infinitesimals of order less than k .

Theorem 9. *If $x \in \bullet\mathbb{R}$ and $k \in \mathbb{N}_{>1}$, then $x^k = 0$ in $\bullet\mathbb{R}$ if and only if ${}^\circ x = 0$ and $\omega(x) < k$.*

Proof: If $x^k = 0$, then taking the standard part map of both sides, we have ${}^\circ(x^k) = ({}^\circ x)^k = 0$ and hence ${}^\circ x = 0$. Moreover $x^k = 0$ means $x_t^k = o(t)$ and hence $\left(\frac{x_t}{t^{1/k}}\right)^k \rightarrow 0$ and $\frac{x_t}{t^{1/k}} \rightarrow 0$. We rewrite this condition using the potential decomposition $x = \sum_{i=1}^k \alpha_i \cdot t^{a_i}$ of x (note that in this way we have $\omega(x) = \frac{1}{a_1}$) obtaining

$$\lim_{t \rightarrow 0^+} \sum_i \alpha_i \cdot t^{a_i - \frac{1}{k}} = 0 = \lim_{t \rightarrow 0^+} t^{a_1 - \frac{1}{k}} \cdot [\alpha_1 + \alpha_2 \cdot t^{a_2 - a_1} + \dots + \alpha_k \cdot t^{a_k - a_1}]$$

But $\alpha_1 + \alpha_2 \cdot t^{a_2 - a_1} + \dots + \alpha_k \cdot t^{a_k - a_1} \rightarrow \alpha_1 \neq 0$, hence we must have that $t^{a_1 - \frac{1}{k}} \rightarrow 0$, and so $a_1 > \frac{1}{k}$, that is $\omega(x) < k$.

Vice versa if ${}^\circ x = 0$ and $\omega(x) < k$, then $x = \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o(t)$, and

$$\lim_{t \rightarrow 0^+} \frac{x_t}{t^{1/k}} = \lim_{t \rightarrow 0^+} \sum_i \alpha_i \cdot t^{a_i - \frac{1}{k}} + \lim_{t \rightarrow 0^+} \frac{o(t)}{t} \cdot t^{1 - \frac{1}{k}}$$

But $t^{1 - \frac{1}{k}} \rightarrow 0$ because $k > 1$ and $t^{a_i - \frac{1}{k}} \rightarrow 0^+$ because $\frac{1}{a_i} \leq \frac{1}{a_1} = \omega(x) < k$ and hence $x^k = 0$ in $\bullet\mathbb{R}$. ■

If we want that in a k -th order infinitesimal neighborhood a smooth function is equal to its k -th Taylor's formula, we need to take infinitesimals which are able to delete the remainder,

that is, such that $h^{k+1} = 0$. The previous theorem permits to extend the definition of the ideal D_k to real number subscripts instead of natural numbers k only.

Definition 10. If $a \in \mathbb{R}_{>0} \cup \{\infty\}$, then

$$D_a := \{x \in {}^\bullet\mathbb{R} \mid {}^\circ x = 0, \omega(x) < a + 1\}$$

Moreover, we will simply denote D_1 by D .

1. If $x = dt_3$, then $\omega(x) = 3$ and $x \in D_3$. More in general $dt_k \in D_a$ if and only if $\omega(dt_k) = k < a + 1$. E.g. $dt_k \in D$ if and only if $1 \leq k < 2$.
2. $D_\infty = \bigcup_a D_a = \{x \in {}^\bullet\mathbb{R} \mid {}^\circ x = 0\}$ is the set of all the infinitesimals of ${}^\bullet\mathbb{R}$.
3. $D_0 = \{0\}$ because the only infinitesimal having order strictly less than 1 is, by definition of order, $x = 0$ (see the Definition 8).

The following theorem gathers several expected properties of the sets D_a and of the order of an infinitesimal $\omega(x)$.

Theorem 11. Let $a, b \in \mathbb{R}_{>0}$ and $x, y \in D_\infty$, then

1. $a \leq b \implies D_a \subseteq D_b$
2. $x \in D_{\omega(x)}$
3. $a \in \mathbb{N} \implies D_a = \{x \in {}^\bullet\mathbb{R} \mid x^{a+1} = 0\}$
4. $x \in D_a \implies x^{[a]+1} = 0$
5. $x \in D_\infty \setminus \{0\}$ and $k = [\omega(x)] \implies x \in D_k \setminus D_{k-1}$
6. $d(x \cdot y) = dx \cdot dy$
7. $x \cdot y \neq 0 \implies \frac{1}{\omega(x \cdot y)} = \frac{1}{\omega(x)} + \frac{1}{\omega(y)}$
8. $x + y \neq 0 \implies \omega(x + y) = \omega(x) \vee \omega(y)$
9. D_a is an ideal

In this statement if $r \in \mathbb{R}$, then $\lceil r \rceil$ is the *ceiling* of the real r , i.e. the unique integer $\lceil r \rceil \in \mathbb{Z}$ such that $\lceil r \rceil - 1 < r \leq \lceil r \rceil$. Moreover if $r, s \in \mathbb{R}$, then $r \vee s := \max(r, s)$.

Property 4. of this theorem cannot be proved substituting the ceiling $\lceil a \rceil$ with the integer part $[a]$. In fact if $a = 1.2$ and $x = dt_{2.1}$, then $\omega(x) = 2.1$ and $[a] + 1 = 2$ so that $x^{[a]+1} = x^2 = dt_{\frac{2.1}{2}} \neq 0$ in $\bullet\mathbb{R}$, whereas $\lceil a \rceil + 1 = 3$ and $x^3 = dt_{\frac{2.1}{3}} = 0$.

Finally let us note the increasing sequence of ideals/neighborhoods of zero:

$$\{0\} = D_0 \subset D = D_1 \subset D_2 \subset \cdots \subset D_k \subset \cdots \subset D_\infty. \quad (16)$$

Because of (16) and of the property $dt_a = 0$ if $a < 1$, we can say that dt is the smallest infinitesimals and dt_2 , dt_3 , etc. are greater infinitesimals; as we will see, this agree to corresponding order properties of these infinitesimals.

VII. PRODUCTS OF POWERS OF NILPOTENT INFINITESIMALS

In this section we will introduce some instruments that will be very useful to decide whether a product of the form $h_1^{i_1} \cdot \dots \cdot h_n^{i_n}$, with $h_k \in D_\infty \setminus \{0\}$, is zero or whether it belongs to some D_k . Generally speaking this problem is not trivial in a ring (e.g. in SDG there is not an effective procedure to decide this problem, see e.g. Lavendhomme [22]) and its solutions will be very useful in the proofs of infinitesimal Taylor's formulas.

Theorem 12. *Let $h_1, \dots, h_n \in D_\infty \setminus \{0\}$ and $i_1, \dots, i_n \in \mathbb{N}$, then*

$$\begin{aligned} 1. \quad h_1^{i_1} \cdot \dots \cdot h_n^{i_n} = 0 & \iff \sum_{k=1}^n \frac{i_k}{\omega(h_k)} > 1 \\ 2. \quad h_1^{i_1} \cdot \dots \cdot h_n^{i_n} \neq 0 & \implies \frac{1}{\omega(h_1^{i_1} \cdot \dots \cdot h_n^{i_n})} = \sum_{k=1}^n \frac{i_k}{\omega(h_k)} \end{aligned}$$

Proof: Let

$$h_k = \sum_{r=1}^{N_k} \alpha_{kr} t^{a_{kr}} \quad (17)$$

be the potential decomposition of h_k for $k = 1, \dots, n$. Then by Definition 7 of potential decomposition and Definition 8 of order, we have $0 < a_{k1} < a_{k2} < \dots < a_{kN_k} \leq 1$ and $j_k := \omega(h_k) = \frac{1}{a_{k1}}$, hence $\frac{1}{j_k} \leq a_{kr}$ for every $r = 1, \dots, N_k$. Therefore from (17), collecting the terms containing t^{1/j_k} we have

$$h_k = t^{1/j_k} \cdot (\alpha_{k1} + \alpha_{k2} t^{a_{k2}-1/j_k} + \dots + \alpha_{kN_k} t^{a_{kN_k}-1/j_k})$$

and hence

$$\begin{aligned} h_1^{i_1} \cdot \dots \cdot h_n^{i_n} &= t^{\frac{i_1}{j_1} + \dots + \frac{i_n}{j_n}} \cdot \left(\alpha_{11} + \alpha_{12} t^{a_{12} - \frac{1}{j_1}} + \dots + \alpha_{1N_1} t^{a_{1N_1} - \frac{1}{j_1}} \right)^{i_1} \cdot \dots \\ &\dots \cdot \left(\alpha_{n1} + \alpha_{n2} t^{a_{n2} - \frac{1}{j_n}} + \dots + \alpha_{nN_n} t^{a_{nN_n} - \frac{1}{j_n}} \right)^{i_n} \end{aligned} \quad (18)$$

Hence if $\sum_k \frac{i_k}{j_k} > 1$ we have that $t^{\frac{i_1}{j_1} + \dots + \frac{i_n}{j_n}} = 0$ in $\bullet\mathbb{R}$, so also $h_1^{i_1} \cdot \dots \cdot h_n^{i_n} = 0$. Vice versa if $h_1^{i_1} \cdot \dots \cdot h_n^{i_n} = 0$, then the right hand side of (18) is a $o(t)$ as $t \rightarrow 0^+$, that is

$$\begin{aligned} t^{\frac{i_1}{j_1} + \dots + \frac{i_n}{j_n} - 1} \cdot \left(\alpha_{11} + \alpha_{12} t^{a_{12} - \frac{1}{j_1}} + \dots + \alpha_{1N_1} t^{a_{1N_1} - \frac{1}{j_1}} \right)^{i_1} \cdot \dots \\ \dots \cdot \left(\alpha_{n1} + \alpha_{n2} t^{a_{n2} - \frac{1}{j_n}} + \dots + \alpha_{nN_n} t^{a_{nN_n} - \frac{1}{j_n}} \right)^{i_n} \rightarrow 0 \end{aligned}$$

But each term $\left(\alpha_{k1} + \alpha_{k2} t^{a_{k2} - \frac{1}{j_k}} + \dots + \alpha_{kN_k} t^{a_{kN_k} - \frac{1}{j_k}} \right)^{i_k} \rightarrow \alpha_k^{i_k} \neq 0$ so, necessarily, we must have $\frac{i_1}{j_1} + \dots + \frac{i_n}{j_n} - 1 > 0$, and this concludes the proof of 1.

To prove 2. it suffices to apply recursively property 7. of Theorem 11. ■

Example 13. $\omega(\mathrm{d}t_{a_1}^{i_1} \cdot \dots \cdot \mathrm{d}t_{a_n}^{i_n})^{-1} = \sum_k \frac{i_k}{\omega(\mathrm{d}t_{a_k})} = \sum_k \frac{i_k}{a_k}$ and $\mathrm{d}t_{a_1}^{i_1} \cdot \dots \cdot \mathrm{d}t_{a_n}^{i_n} = 0$ if and only if $\sum_k \frac{i_k}{a_k} > 1$, so e.g. $\mathrm{d}t \cdot h = 0$ for every $h \in D_\infty$.

The following corollary gives a necessary and sufficient condition to have $h_1^{i_1} \cdot \dots \cdot h_n^{i_n} \in D_p \setminus \{0\}$.

Corollary 14. *In the hypotheses of the previous Theorem 12 let $p \in \mathbb{R}_{>0}$, then we have*

$$h_1^{i_1} \cdot \dots \cdot h_n^{i_n} \in D_p \setminus \{0\} \quad \Longleftrightarrow \quad \frac{1}{p+1} < \sum_{k=1}^n \frac{i_k}{\omega(h_k)} \leq 1$$

Let $h, k \in D$; because in this case $\sum_k \frac{i_k}{j_k+1} = \frac{1}{2} + \frac{1}{2} = 1$ we always have

$$h \cdot k = 0. \quad (19)$$

This is a great conceptual difference between Fermat reals and the ring of SDG, where, not necessarily, the product of two first order infinitesimal is zero. The consequences of this property of Fermat reals arrive very deeply in the development of the theory of Fermat reals, forcing us, e.g., to develop several new concepts if we want to generalize the derivation formula (4) to functions defined on infinitesimal domains, like $f : D \longrightarrow \bullet\mathbb{R}$ (see Giordano [16]). We only mention here that looking at the simple Definition 4, the equality (19) has

an intuitively clear meaning, and it is to preserve this intuition that we keep this equality instead of changing completely the theory toward a less intuitive one.

Let us note explicitly that the possibility to prove these results about products of powers of nilpotent infinitesimals is essentially tied with the choice of little-oh polynomials in the definition of the equivalence relation \sim in Definition 2. Equally effective and useful results are not provable for the more general family of nilpotent functions (see e.g. Giordano [15]).

VIII. IDENTITY PRINCIPLE FOR POLYNOMIALS AND INVERTIBLE FERMAT REALS

In this section we want to prove that if a polynomial $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ of $\bullet\mathbb{R}$ is identically zero, then $a_k = 0$ for all $k = 0, \dots, n$. To prove this conclusion, it suffices to mean “identically zero” as “equal to zero for every x belonging to the extension of an open subset of \mathbb{R} ”. Therefore we firstly define what this extension is.

Definition 15. If U is an open subset of \mathbb{R}^n , then $\bullet U := \{x \in \bullet\mathbb{R}^n \mid x \in U\}$. Here with the symbol $\bullet\mathbb{R}^n$ we mean $\bullet\mathbb{R}^n := \bullet\mathbb{R} \times \dots \times \bullet\mathbb{R}$.

The identity principle for polynomials can now be stated in the following way and proved in standard manner using Vandermonde matrices.

Theorem 16. Let $a_0, \dots, a_n \in \bullet\mathbb{R}$ and U be an open neighborhood of 0 in \mathbb{R} such that

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0 \quad \text{in } \bullet\mathbb{R} \quad \forall x \in \bullet U \quad (20)$$

Then

$$a_0 = a_1 = \dots = a_n = 0 \quad \text{in } \bullet\mathbb{R}$$

Now, we want to see more formally that to prove (3) we cannot embed the reals \mathbb{R} into a field but only into a ring, necessarily containing nilpotent element. In fact, applying (3) to the function $f(h) = h^2$ for $h \in D$, where $D \subseteq \bullet\mathbb{R}$ is a given subset of $\bullet\mathbb{R}$, we have

$$f(h) = h^2 = f(0) + h \cdot f'(0) = 0 \quad \forall h \in D.$$

Where we have supposed the preservation of the equality $f'(0) = 0$ from \mathbb{R} to $\bullet\mathbb{R}$. In other words, if D and $f(h) = h^2$ verify (3), then necessarily each element $h \in D$ must be a new type of number whose square is zero.

Because we cannot have property (3) and a field at the same time, we need a sufficiently good family of cancellation laws as substitutes. The simplest one of them is also useful to prove the uniqueness of (4):

Theorem 17. *If $x \in \bullet\mathbb{R}$ is a Fermat real and $r, s \in \mathbb{R}$ are standard real numbers, then*

$$x \cdot r = x \cdot s \text{ in } \bullet\mathbb{R} \quad \text{and} \quad x \neq 0 \quad \implies \quad r = s$$

Proof: From the Definition 4 of equality in $\bullet\mathbb{R}$ and from $x \cdot r = x \cdot s$ we have

$$\lim_{t \rightarrow 0^+} \frac{x_t \cdot (r - s)}{t} = 0.$$

But if we had $r \neq s$ this would implies $\lim_{t \rightarrow 0^+} \frac{x_t}{t} = 0$, that is $x = 0$ in $\bullet\mathbb{R}$ and this contradicts the hypothesis $x \neq 0$. ■

The last result of this section takes its ideas from similar situations of formal power series and gives also a formula to compute the inverse of an invertible Fermat real.

Theorem 18. *Let $x = {}^\circ x + \sum_{i=1}^n {}^\circ x_i \cdot dt_{a_i}$ be the decomposition of a Fermat real $x \in \bullet\mathbb{R}$. Then x is invertible if and only if ${}^\circ x \neq 0$, and in this case*

$$\frac{1}{x} = \frac{1}{{}^\circ x} \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot \left(\sum_{i=1}^n \frac{{}^\circ x_i}{{}^\circ x} \cdot dt_{a_i} \right)^j \quad (21)$$

In the formula (21) we have to note that the series is actually a finite sum because any dt_{a_i} is nilpotent, e.g. $(1 + dt_2)^{-1} = 1 - dt_2 + dt_2^2 - dt_2^3 + \dots = 1 - dt_2 + dt$ because $dt_2^3 = 0$.

Proof: If $x \cdot y = 1$ for some $y \in \bullet\mathbb{R}$, then, taking the standard parts of each side we have ${}^\circ x \cdot {}^\circ y = 1$ and hence ${}^\circ x \neq 0$. Vice versa let $y := {}^\circ x^{-1} \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot (\sum_i \frac{{}^\circ x_i}{{}^\circ x} dt_{a_i})^j$ and $h := x - {}^\circ x = \sum_i {}^\circ x_i dt_{a_i} \in D_\infty$ so that we can also write

$$y = {}^\circ x^{-1} \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot \frac{h^j}{{}^\circ x^j}$$

But $h \in \bullet\mathbb{R}$ is a little-oh polynomial with $h(0) = 0$, so it is also continuous, hence for a sufficiently small $\delta > 0$ we have

$$\forall t \in (-\delta, \delta) : \quad \left| \frac{h_t}{{}^\circ x} \right| < 1.$$

Therefore

$$\forall t \in (-\delta, \delta) : \quad y_t = \frac{1}{{}^\circ x} \cdot \left(1 + \frac{h_t}{{}^\circ x} \right)^{-1} = \frac{1}{{}^\circ x + h_t} = \frac{1}{x_t}$$

From this equality and from Definition 4 it follows $x \cdot y = 1$ in $\bullet\mathbb{R}$. ■

IX. THE DERIVATION FORMULA

In this section we want to give a proof of (4) because it has been the principal motivation for the construction of the ring of Fermat reals $\bullet\mathbb{R}$. Anyhow, before considering the proof of the derivation formula, we have to extend a given smooth function $f : \mathbb{R} \longrightarrow \mathbb{R}$ to a certain function $\bullet f : \bullet\mathbb{R} \longrightarrow \bullet\mathbb{R}$.

Definition 19. Let A be an open subset of \mathbb{R}^n , $f : A \longrightarrow \mathbb{R}$ a smooth function and $x \in \bullet A$ then we define

$$\bullet f(x) := f \circ x.$$

This definition is correct because we have seen that little-oh polynomials are preserved by smooth functions, and because the function f is locally Lipschitz, so

$$\left| \frac{f(x_t) - f(y_t)}{t} \right| \leq K \cdot \left| \frac{x_t - y_t}{t} \right| \quad \forall t \in (-\delta, \delta)$$

for a sufficiently small δ and some constant K , and hence if $x = y$ in $\bullet\mathbb{R}$, then also $\bullet f(x) = \bullet f(y)$ in $\bullet\mathbb{R}$.

The function $\bullet f$ is an extension of f , that is

$$\bullet f(r) = f(r) \quad \text{in } \bullet\mathbb{R} \quad \forall r \in \mathbb{R},$$

as it follows directly from the definition of equality in $\bullet\mathbb{R}$ (i.e. Definition 4), thus we can still use the symbol $f(x)$ both for $x \in \bullet\mathbb{R}$ and $x \in \mathbb{R}$ without confusion. After the introduction of the extension of smooth functions, we can also state the following useful *elementary transfer theorem* for equalities, whose proof follows directly from the previous definitions:

Theorem 20. Let A be an open subset of \mathbb{R}^n , and $\tau, \sigma : A \longrightarrow \mathbb{R}$ be smooth functions. Then it results

$$\forall x \in \bullet A : \bullet \tau(x) = \bullet \sigma(x)$$

iff

$$\forall r \in A : \tau(r) = \sigma(r).$$

Now we will prove the derivation formula (4).

Theorem 21. Let A be an open set in \mathbb{R} , $x \in A$ and $f : A \longrightarrow \mathbb{R}$ a smooth function, then

$$\exists! m \in \mathbb{R} \quad \forall h \in D : f(x + h) = f(x) + h \cdot m. \tag{22}$$

In this case we have $m = f'(x)$, where $f'(x)$ is the usual derivative of f at x .

Proof: Uniqueness follows from the previous cancellation law Theorem 17, indeed if $m_1 \in \mathbb{R}$ and $m_2 \in \mathbb{R}$ both verify (22), then $h \cdot m_1 = h \cdot m_2$ for every $h \in D$. But there exists a non zero first order infinitesimal, e.g. $dt \in D$, so from Theorem (17) it follows $m_1 = m_2$.

To prove the existence part, take $h \in D$, so that $h^2 = 0$ in $\bullet\mathbb{R}$, i.e. $h_t^2 = o(t)$ for $t \rightarrow 0^+$. But f is smooth, hence from its second order Taylor's formula we have

$$f(x + h_t) = f(x) + h_t \cdot f'(x) + \frac{h_t^2}{2} \cdot f''(x) + o(h_t^2)$$

But

$$\frac{o(h_t^2)}{t} = \frac{o(h_t^2)}{h_t^2} \cdot \frac{h_t^2}{t} \rightarrow 0 \quad \text{for } t \rightarrow 0^+$$

so

$$\frac{h_t^2}{2} \cdot f''(x) + o(h_t^2) = o_1(t) \quad \text{for } t \rightarrow 0^+$$

and we can write

$$f(x + h_t) = f(x) + h_t \cdot f'(x) + o_1(t) \quad \text{for } t \rightarrow 0^+$$

that is

$$f(x + h) = f(x) + h \cdot f'(x) \quad \text{in } \bullet\mathbb{R}$$

and this proves the existence part because $f'(x) \in \mathbb{R}$. ■

For example $e^h = 1 + h$, $\sin(h) = h$ and $\cos(h) = 1$ for every $h \in D$.

Analogously we can prove the following infinitesimal Taylor's formula.

Lemma 22. *Let A be an open set in \mathbb{R}^d , $x \in A$, $n \in \mathbb{N}_{>0}$ and $f : A \longrightarrow \mathbb{R}$ a smooth function, then*

$$\forall h \in D_n^d : f(x + h) = \sum_{\substack{j \in \mathbb{N}^d \\ |j| \leq n}} \frac{h^j}{j!} \cdot \frac{\partial^{|j|} f}{\partial x^j}(x)$$

For example $\sin(h) = h - \frac{h^3}{6}$ if $h \in D_3$ so that $h^4 = 0$.

It is possible to generalize several results of the present work to functions of class \mathcal{C}^n only, instead of smooth ones. However it is an explicit purpose of this work to simplify statements of results, definitions and notations, even if, as a result of this searching for simplicity, its applicability will only hold for a more restricted class of functions. Some more general results, stated for \mathcal{C}^n functions, but less simple can be found in Giordano [15].

Note that $m = f'(x) \in \mathbb{R}$, i.e. the slope is a standard real number, and that we can use the previous formula with standard real numbers x only, and not with a generic $x \in \bullet\mathbb{R}$, but we shall remove this limitation in subsequent works (see also Giordano [16]).

If we apply this theorem to the smooth function $p(r) := \int_x^{x+r} f(t) dt$, for f smooth, then we immediately obtain the following result frequently used in several informal calculations:

Corollary 23. *Let A be open in \mathbb{R} , $x \in A$ and $f : A \longrightarrow \mathbb{R}$ smooth. Then*

$$\forall h \in D : \int_x^{x+h} f(t) dt = h \cdot f(x).$$

Moreover $f(x) \in \mathbb{R}$ is uniquely determined by this equality.

X. NILPOTENT INFINITESIMALS AND ORDER PROPERTIES

Like in other disciplines, also in mathematics the layout of a work reflects the personal philosophical ideas of the authors. In particular the present work is based on the idea that a good mathematical theory is able to construct a good dialectic between formal properties, proved in the theory, and their informal interpretations. The dialectic has to be, as far as possible, in both directions: theorems proved in the theory should have a clear and useful intuitive interpretation and, on the other hand, the intuition corresponding to the theory has to be able to suggest true sentences, i.e. conjectures or sketch of proofs that can then be converted into rigorous proofs.

In a theory of new numbers, like the present one about Fermat reals, the introduction of an order relation can be a hard test of the excellence of this dialectic between formal properties and their informal interpretations. Indeed if we introduce a new ring of numbers (like $\bullet\mathbb{R}$) extending the real field \mathbb{R} , we want that the new order relation, defined on the new ring, will extend the standard one on \mathbb{R} . This extension naturally leads to the wish of finding a geometrical representation of the new numbers, according to the above principle of having a good formal/informal dialectic.

We want to start this section showing that in our setting there is a strong connection between some order properties and some algebraic properties. In particular, we will show that it is not possible to have good order properties and at the same time a uniqueness without limitations in the derivation formula. In the following theorem we can see that the property $h \cdot k = 0$ is a general consequence if we suppose to have a total order on D .

Theorem 24. *Let (R, \leq) be a generic ordered ring and $D \subseteq R$ a subset of this ring, such that*

1. $0 \in D$
2. $\forall h \in D : h^2 = 0 \text{ and } -h \in D$
3. (D, \leq) is a total order

then $h \cdot k = 0$ for every $h, k \in D$.

This theorem implies that if we want a total order in our theory of infinitesimal numbers, and if in this theory we consider $D = \{h \mid h^2 = 0\}$, then we must accept that the product of any two elements of D must be zero. For example, if we think that a geometric representation of infinitesimals is not possible if we do not have, at least, the trichotomy law, then in this theory we must also have that the product of two first order infinitesimals is zero.

Proof: Let $h, k \in D$ be two elements of the subset D . By hypotheses $0, -h, -k \in D$, hence all these elements are comparable with respect to the order relation \leq , because, by hypotheses this relation is total in D . E.g.

$$h \leq k \quad \text{or} \quad k \leq h$$

We will consider only the case $h \leq k$, because analogously we can deal with the case $k \leq h$, simply exchanging everywhere h with k and vice versa.

First sub-case: $k \geq 0$. By multiplying both sides of $h \leq k$ by $k \geq 0$ we obtain

$$hk \leq k^2 \tag{23}$$

If $h \geq 0$ then, multiplying by $k \geq 0$ we have $0 \leq hk$, so from (23) we have $0 \leq hk \leq k^2 = 0$, and hence $hk = 0$.

If $h \leq 0$ then, multiplying by $k \geq 0$ we have

$$hk \leq 0 \tag{24}$$

If, furthermore, $h \geq -k$, then multiplying by $k \geq 0$ we have $hk \geq -k^2$, hence from (24) $0 \geq hk \geq -k^2 = 0$, hence $hk = 0$.

If, otherwise, $h \leq -k$, then multiplying by $-h \geq 0$ we have $-h^2 = 0 \leq hk \leq 0$ from (24), hence $hk = 0$. This concludes the discussion of the case $k \geq 0$.

Second sub-case: $k \leq 0$. In this case we have $h \leq k \leq 0$. Multiplying both inequalities by $h \leq 0$ we obtain $h^2 = 0 \geq hk \geq 0$ and hence $hk = 0$. ■

So, the trichotomy law is incompatible with the uniqueness in a possible derivation formula like

$$\exists! m \in R : \forall h \in D : f(h) = f(0) + h \cdot m \quad (25)$$

framed in the ring R of Theorem 24. In fact, if $a, b \in D$ are two elements of the subset $D \subseteq R$, then both a and b play the role of $m \in R$ in (25) for the linear function

$$f : h \in D \mapsto h \cdot a = 0 \in R$$

So, if the derivation formula (25) applies to linear functions (or less, to constant functions), the uniqueness part of this formula cannot hold in the ring R .

In the next section we will introduce a natural and meaningful total order relation on $\bullet\mathbb{R}$. Therefore, the previous Theorem 24 strongly motivate that for the ring of Fermat reals $\bullet\mathbb{R}$ we must have that the product of two first order infinitesimals must be zero and hence, that for the derivation formula in $\bullet\mathbb{R}$ the uniqueness cannot hold in its strongest form. Since we will also see that the order relation permits to have a geometric representation of Fermat reals, we can summarize the conclusions of this section saying that the uniqueness in the derivation formula is incompatible with a natural geometric interpretation of Fermat reals and hence with a good dialectic between formal properties and informal interpretations in this theory.

XI. ORDER RELATION

From the previous sections one can draw the conclusion that the ring of Fermat reals $\bullet\mathbb{R}$ is essentially “the little-oh” calculus. But, on the other hand the Fermat reals give us more flexibility than this calculus: working with $\bullet\mathbb{R}$ we do not have to bother ourselves with remainders made of “little-oh”, but we can neglect them and use the powerful algebraic calculus with nilpotent infinitesimals. But thinking the elements of $\bullet\mathbb{R}$ as new numbers, and not simply as “little-oh functions”, permits to treat them in a different and new way, for example to define on them an order relation with a clear geometrical interpretation.

First of all, let us introduce the useful notation

$$\forall^0 t \geq 0 : \mathcal{P}(t)$$

and we will read the quantifier $\forall^0 t \geq 0$ saying “for every $t \geq 0$ (sufficiently) small”, to indicate that the property $\mathcal{P}(t)$ is true for all t in some right neighborhood of $t = 0$ (recall that, by Definition 2, our little-oh polynomials are always defined on $\mathbb{R}_{\geq 0}$), i.e.

$$\exists \delta > 0 : \forall t \in [0, \delta) : \mathcal{P}(t).$$

The first heuristic idea to define an order relation is the following

$$x \leq y \iff x - y \leq 0 \iff \exists z : z = 0 \text{ in } \bullet\mathbb{R} \quad \text{and} \quad x - y \leq z$$

More formally:

Definition 25. Let $x, y \in \bullet\mathbb{R}$, then we say

$$x \leq y$$

iff we can find $z \in \bullet\mathbb{R}$ such that $z = 0$ in $\bullet\mathbb{R}$ and

$$\forall^0 t \geq 0 : x_t \leq y_t + z_t$$

Recall that $z = 0$ in $\bullet\mathbb{R}$ is equivalent to $z_t = o(t)$ for $t \rightarrow 0^+$. It is immediate to see that we can equivalently define $x \leq y$ if and only if we can find $x' = x$ and $y' = y$ in $\bullet\mathbb{R}$ such that $x_t \leq y_t$ for every t sufficiently small. From this it also follows that the relation \leq is well defined on $\bullet\mathbb{R}$, i.e. if $x' = x$ and $y' = y$ in $\bullet\mathbb{R}$ and $x \leq y$, then $x' \leq y'$ (recall that, to simplify the notations, we do not use equivalence classes as elements of $\bullet\mathbb{R}$ but directly little-oh polynomials). As usual we will use the notation $x < y$ for $x \leq y$ and $x \neq y$.

Theorem 26. *The relation \leq is an order, i.e. is reflexive, transitive and anti-symmetric; it extends the order relation of \mathbb{R} and with it $(\bullet\mathbb{R}, \leq)$ is an ordered ring. Finally the following sentences are equivalent:*

1. $h \in D_\infty$, i.e. h is an infinitesimal

2. $\forall r \in \mathbb{R}_{>0} : -r < h < r$

Hence an infinitesimal can be thought of as a number with standard part zero, or as a number smaller than every standard positive real number and greater than every standard negative real number.

Proof: We only prove the property

$$x \leq y \quad \text{and} \quad w \geq 0 \quad \implies \quad x \cdot w \leq y \cdot w,$$

the others being a simple consequence of our Definition 25. Let us suppose that

$$\begin{aligned} x_t &\leq y_t + z_t \quad \forall^0 t \geq 0 \\ w_t &\geq z'_t \quad \forall^0 t \geq 0 \end{aligned} \tag{26}$$

then $w_t - z'_t \geq 0$ for every t small and hence from (26)

$$x_t \cdot (w_t - z'_t) \leq y_t \cdot (w_t - z'_t) + z_t \cdot (w_t - z'_t) \quad \forall^0 t \geq 0$$

from which it follows

$$x_t \cdot w_t \leq y_t \cdot w_t + (-x_t z'_t - y_t z'_t + z_t w_t - z_t z'_t) \quad \forall^0 t \geq 0$$

But $-xz' - yz' + zw - zz' = 0$ in ${}^\bullet\mathbb{R}$ because $z = 0$ and $z' = 0$ and hence the conclusion follows. ■

Example. We have e.g. $dt > 0$ and $dt_2 - 3dt > 0$ because for $t \geq 0$ sufficiently small $t^{1/2} > 3t$ and hence

$$t^{1/2} - 3t > 0 \quad \forall^0 t \geq 0.$$

From examples like these ones we can guess that our little-oh polynomials are always locally comparable with respect to pointwise order relation, and this is the first step to prove that for our order relation the trichotomy law holds. In the following statement we will use the notation $\forall^0 t > 0 : \mathcal{P}(t)$, that naturally means

$$\forall^0 t \geq 0 : t \neq 0 \quad \implies \quad \mathcal{P}(t)$$

where $\mathcal{P}(t)$ is a generic property depending on t .

Lemma 27. *Let $x, y \in {}^\bullet\mathbb{R}$, then*

$$1. \quad {}^\circ x < {}^\circ y \implies \forall^0 t \geq 0 : x_t < y_t$$

2. If ${}^\circ x = {}^\circ y$, then

$$(\forall^0 t > 0 : x_t < y_t) \quad \text{or} \quad (\forall^0 t > 0 : x_t > y_t) \quad \text{or} \quad (x = y \text{ in } {}^\bullet \mathbb{R})$$

Proof:

1.) Let us suppose that ${}^\circ x < {}^\circ y$, then the continuous function $t \geq 0 \mapsto y_t - x_t \in \mathbb{R}$ assumes the value $y_0 - x_0 > 0$ hence is locally positive, i.e.

$$\forall^0 t \geq 0 : x_t < y_t$$

2.) Now let us suppose that ${}^\circ x = {}^\circ y$, and introduce a notation for the potential decompositions of x and y (see Definition 7). From the definition of equality in ${}^\bullet \mathbb{R}$, we can always write

$$\begin{aligned} x_t &= {}^\circ x + \sum_{i=1}^N \alpha_i \cdot t^{a_i} + z_t \quad \forall t \geq 0 \\ y_t &= {}^\circ y + \sum_{j=1}^M \beta_j \cdot t^{b_j} + w_t \quad \forall t \geq 0 \end{aligned}$$

where $x = {}^\circ x + \sum_{i=1}^N \alpha_i \cdot t^{a_i}$ and $y = {}^\circ y + \sum_{j=1}^M \beta_j \cdot t^{b_j}$ are the potential decompositions of x and y (hence $0 < \alpha_i < \alpha_{i+1} \leq 1$ and $0 < \beta_j < \beta_{j+1} \leq 1$), whereas w and z are little-oh polynomials such that $z_t = o(t)$ and $w_t = o(t)$ for $t \rightarrow 0^+$.

Case: $a_1 < b_1$ In this case the least power in the two decompositions is $\alpha_1 \cdot t^{a_1}$, and hence we expect that the second alternative of the conclusion is the true one if $\alpha_1 > 0$, otherwise the first alternative will be the true one if $\alpha_1 < 0$ (recall that always $\alpha_i \neq 0$ in a decomposition). Indeed, let us analyze, for $t > 0$, the condition $x_t < y_t$: the following formulae are all equivalent to it

$$\begin{aligned} \sum_{i=1}^N \alpha_i \cdot t^{a_i} &< \sum_{j=1}^M \beta_j \cdot t^{b_j} + w_t - z_t \\ t^{a_1} \cdot \left[\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} \right] &< t^{a_1} \cdot \left[\sum_{j=1}^M \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1} \right] \\ \alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} &< \sum_{j=1}^M \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1}. \end{aligned}$$

Therefore, let us consider the function

$$f(t) := \sum_{j=1}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1} - \alpha_1 - \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} \quad \forall t \geq 0$$

We can write

$$(w_t - z_t) \cdot t^{-a_1} = \frac{w_t - z_t}{t} \cdot t^{1-a_1}$$

and $\frac{w_t - z_t}{t} \rightarrow 0$ as $t \rightarrow 0^+$ because $w_t = o(t)$ and $z_t = o(t)$. Furthermore, $a_1 \leq 1$ hence t^{1-a_1} is bounded in a right neighborhood of $t = 0$. Therefore, $(w_t - z_t) \cdot t^{-a_1} \rightarrow 0$ and the function f is continuous at $t = 0$ too, because $a_i < a_i$ and $a_1 < b_1 < b_j$. By continuity, the function f is locally strictly positive if and only if $f(0) = -\alpha_1 > 0$, hence

$$(\forall^0 t > 0 : x_t < y_t) \iff \alpha_1 < 0$$

$$(\forall^0 t > 0 : x_t > y_t) \iff \alpha_1 > 0$$

Case: $a_1 > b_1$ We can argue in an analogous way with b_1 and β_1 instead of a_1 and α_1 .

Case: $a_1 = b_1$ We shall exploit the same idea used above and analyze the condition $x_t < y_t$.

The following are equivalent ways to express this condition

$$\begin{aligned} t^{a_1} \cdot \left[\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} \right] &< t^{a_1} \cdot \left[\beta_1 + \sum_{j=2}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1} \right] \\ \alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} &< \beta_1 + \sum_{j=2}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1} \end{aligned}$$

Hence, exactly as we have demonstrated above, we can state that

$$\alpha_1 < \beta_1 \implies \forall^0 t > 0 : x_t < y_t$$

$$\alpha_1 > \beta_1 \implies \forall^0 t > 0 : x_t > y_t$$

Otherwise $\alpha_1 = \beta_1$ and we can restart with the same reasoning using $a_2, b_2, \alpha_2, \beta_2$, etc. If $N = M$, the number of addends in the decompositions, using this procedure we can prove that

$$\forall t \geq 0 : x_t = y_t + w_t - z_t,$$

that is $x = y$ in $\bullet\mathbb{R}$.

It remains to consider the case, e.g., $N < M$. In this hypotheses, using the previous procedure we would arrive at the following analysis of the condition $x_t < y_t$:

$$0 < \sum_{j>N} \beta_j \cdot t^{b_j} + w_t - z_t$$

$$0 < t^{b_{N+1}} \cdot \left[\beta_{N+1} + \sum_{j>N+1} \beta_j \cdot t^{b_j - b_{N+1}} + (w_t - z_t) \cdot t^{-b_{N+1}} \right]$$

$$0 < \beta_{N+1} + \sum_{j>N+1} \beta_j \cdot t^{b_j - b_{N+1}} + (w_t - z_t) \cdot t^{-b_{N+1}}$$

Hence

$$\beta_{N+1} > 0 \implies \forall^0 t > 0 : x_t < y_t$$

$$\beta_{N+1} < 0 \implies \forall^0 t > 0 : x_t > y_t$$

■

This lemma can be used to find an equivalent formulation of the order relation.

Theorem 28. *Let $x, y \in {}^\bullet\mathbb{R}$, then*

1. $x \leq y \iff (\forall^0 t > 0 : x_t < y_t) \quad \text{or} \quad (x = y \text{ in } {}^\bullet\mathbb{R})$
2. $x < y \iff (\forall^0 t > 0 : x_t < y_t) \quad \text{and} \quad (x \neq y \text{ in } {}^\bullet\mathbb{R})$

Proof:

1.) \Rightarrow If ${}^\circ x < {}^\circ y$ then, from the previous Lemma 27 we can derive that the first alternative is true. If ${}^\circ x = {}^\circ y$, then from Lemma 27 we have

$$(\forall^0 t > 0 : x_t < y_t) \quad \text{or} \quad (x = y \text{ in } {}^\bullet\mathbb{R}) \quad \text{or} \quad (\forall^0 t > 0 : x_t > y_t) \quad (27)$$

In the first two cases we have the conclusion. In the third case, from $x \leq y$ we obtain

$$\forall^0 t \geq 0 : x_t \leq y_t + z_t \quad (28)$$

with $z_t = o(t)$. Hence from the third alternative of (27) we have

$$0 < x_t - y_t \leq z_t \quad \forall^0 t > 0$$

and hence $\lim_{t \rightarrow 0^+} \frac{x_t - y_t}{t} = 0$, i.e. $x = y$ in ${}^\bullet\mathbb{R}$.

- 1.) \Leftarrow This follows immediately from the reflexive property of \leq or from the Definition 25.
- 2.) \Rightarrow From $x < y$ we have $x \leq y$ and $x \neq y$, so the conclusion follows from the previous 1.
- 2.) \Leftarrow From $\forall^0 t > 0 : x_t < y_t$ and from 1. it follows $x \leq y$ and hence $x < y$ from the hypotheses $x \neq y$.

■

Now we can prove that our order is total

Corollary 29. *Let $x, y \in {}^\bullet\mathbb{R}$, then in ${}^\bullet\mathbb{R}$ we have*

$$1. \ x \leq y \quad \text{or} \quad y \leq x \quad \text{or} \quad x = y$$

$$2. \ x < y \quad \text{or} \quad y < x \quad \text{or} \quad x = y$$

Proof:

1.) If ${}^\circ x < {}^\circ y$, then from Lemma 27 we have $x_t < y_t$ for $t \geq 0$ sufficiently small. Hence from Theorem 28 we have $x \leq y$. We can argue in the same way if ${}^\circ x > {}^\circ y$. Also the case ${}^\circ x = {}^\circ y$ can be handled in the same way using 2. of Lemma 27.

2.) This part is a general consequence of the previous one. ■

From the proof of Lemma 27 and from Theorem 28 we can deduce the following

Theorem 30. *Let $x, y \in {}^\bullet\mathbb{R}$. If ${}^\circ x \neq {}^\circ y$, then*

$$x < y \iff {}^\circ x < {}^\circ y$$

Otherwise, if ${}^\circ x = {}^\circ y$, then

$$1. \text{ If } \omega(x) > \omega(y), \text{ then } x > y \text{ iff } {}^\circ x_1 > 0$$

$$2. \text{ If } \omega(x) = \omega(y), \text{ then}$$

$${}^\circ x_1 > {}^\circ y_1 \implies x > y$$

$${}^\circ x_1 < {}^\circ y_1 \implies x < y$$

Example. The previous Theorem gives an effective criterion to decide whether $x < y$ or not. Indeed, if the two standard parts are different, then the order relation can be decided on the basis of these standard parts only. E.g. $2 + dt_2 > 3dt$ and $1 + dt_2 < 3 + dt$.

Otherwise, if the standard parts are equal, we firstly have to look at the order and at the first standard parts, i.e. ${}^\circ x_1$ and ${}^\circ y_1$, which are the coefficients of the biggest infinitesimals in the decompositions of x and y . E.g. $3dt_2 > 5dt$, and $dt_2 > a dt$ for every $a \in \mathbb{R}$, and $dt < dt_2 < dt_3 < \dots < dt_k$ for every $k > 3$, and $dt_k > 0$.

If the orders are equal we have to compare the first standard parts. E.g. $3dt_5 > 2dt_5$.

The other cases fall within the previous ones, because of the properties of the ordered ring ${}^\bullet\mathbb{R}$. E.g. we have that $dt_5 - 2dt_3 + 3dt < dt_5 - 2dt_3 + dt_{3/2}$ if and only if $3dt < dt_{3/2}$, which is true because $\omega(dt) = 1 < \omega(dt_{3/2}) = \frac{3}{2}$. Finally $dt_5 - 2dt_3 + 3dt > dt_5 - 2dt_3 - dt$ because $3dt > -dt$.

XII. ABSOLUTE VALUE, POWERS AND LOGARITHMS

Having a total order we can define the absolute value in the usual way, and, exactly like for the real field \mathbb{R} , we can prove the usual properties of the absolute value. Moreover, also the following cancellation law is provable.

Theorem 31. *Let $h \in \bullet\mathbb{R} \setminus \{0\}$ and $r, s \in \mathbb{R}$, then*

$$|h| \cdot r \leq |h| \cdot s \implies r \leq s$$

Proof: In fact if $|h| \cdot r \leq |h| \cdot s$ then from Theorem 28 we obtain that either

$$\forall^0 t > 0 : |h_t| \cdot r \leq |h_t| \cdot s \tag{29}$$

or $|h| \cdot r = |h| \cdot s$. But $h \neq 0$ so

$$(\forall^0 t > 0 : h_t > 0) \quad \text{or} \quad (\forall^0 t > 0 : h_t < 0)$$

hence we can always find a $\bar{t} > 0$ such that $|h_{\bar{t}}| \neq 0$ and to which (29) is applicable.

Therefore, in the first case we must have $r \leq s$. In the second one we have

$$|h| \cdot r = |h| \cdot s$$

but $h \neq 0$, hence $|h| \neq 0$ and so the conclusion follows from Theorem 17. ■

Due to the presence of nilpotent elements in $\bullet\mathbb{R}$, we cannot define powers x^y and logarithms $\log_x y$ without any limitation. E.g. we cannot define the square root having the usual properties, like

$$x \in \bullet\mathbb{R} \implies \sqrt{x} \in \bullet\mathbb{R} \tag{30}$$

$$x = y \text{ in } \bullet\mathbb{R} \implies \sqrt{x} = \sqrt{y} \text{ in } \bullet\mathbb{R} \tag{31}$$

$$\sqrt{x^2} = |x|$$

because they are incompatible with the existence of $h \in D$ such that $h^2 = 0$, but $h \neq 0$. Indeed, the general property stated in the Subsection IV permits to obtain a property like (30) (i.e. the closure of $\bullet\mathbb{R}$ with respect to a given operation) only for smooth functions. Moreover, the Definition 19 states that to obtain a well defined operation we need a locally Lipschitz function. For these reasons, we will limit x^y to $x > 0$ and x invertible only, and $\log_x y$ to $x, y > 0$ and both x, y invertible.

Definition 32. Let $x, y \in {}^\bullet\mathbb{R}$, with x strictly positive and invertible, then

1. $x^y := [t \geq 0 \mapsto x_t^y]_{= \text{in } {}^\bullet\mathbb{R}}$
2. If $y > 0$ and y is invertible, then $\log_x y := [t \geq 0 \mapsto \log_{x_t} y_t]_{= \text{in } {}^\bullet\mathbb{R}}$

Because of Theorem 28 from $x > 0$ we have

$$\forall^0 t > 0 : x_t > 0$$

so that, exactly as we proved in Subsection IV and in Definition 19, the previous operations are well defined in ${}^\bullet\mathbb{R}$ because ${}^\circ x \neq 0 \neq {}^\circ y$. From the elementary transfer theorem 20 the usual properties follow. To prove the usual monotonicity properties, it suffices to use Theorem 28.

Finally, it can be useful to state here the *elementary transfer theorem for inequalities*, whose proof follows immediately from the definition of \leq and from Theorem 28:

Theorem 33. Let A be an open subset of \mathbb{R}^n , and $\tau, \sigma : A \longrightarrow \mathbb{R}$ be smooth functions. Then

$$\forall x \in {}^\bullet A : {}^\bullet \tau(x) \leq {}^\bullet \sigma(x)$$

iff

$$\forall r \in A : \tau(r) \leq \sigma(r).$$

XIII. GEOMETRICAL REPRESENTATION OF FERMAT REALS

At the beginning of this article we argued that one of the conducting idea in the construction of Fermat reals is to maintain always a clear intuitive meaning. More precisely, we always tried, and we will always try, to keep a good dialectic between provable formal properties and their intuitive meaning. In this direction we can see the possibility to find a geometrical representation of Fermat reals.

The idea is that to any Fermat real $x \in {}^\bullet\mathbb{R}$ we can associate the function

$$t \in \mathbb{R}_{\geq 0} \mapsto {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} \in \mathbb{R} \quad (32)$$

where N is, of course, the number of addends in the decomposition of x . Therefore, a geometric representation of this function is also a geometric representation of the number

x , because different Fermat reals have different decompositions, see Theorem 5. Finally, we can guess that, because the notion of equality in $\bullet\mathbb{R}$ depends only on the germ generated by each little-oh polynomial (see Definition 4), we can represent each $x \in \bullet\mathbb{R}$ with only the first small part of the function (32).

Definition 34. If $x \in \bullet\mathbb{R}$ and $\delta \in \mathbb{R}_{>0}$, then

$$\text{graph}_\delta(x) := \left\{ \left({}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)}, t \right) \mid 0 \leq t < \delta \right\}$$

where N is the number of addends in the decomposition of x .

Note that the value of the function are placed in the abscissa position, so that the correct representation of $\text{graph}_\delta(x)$ is given by the figure 1.

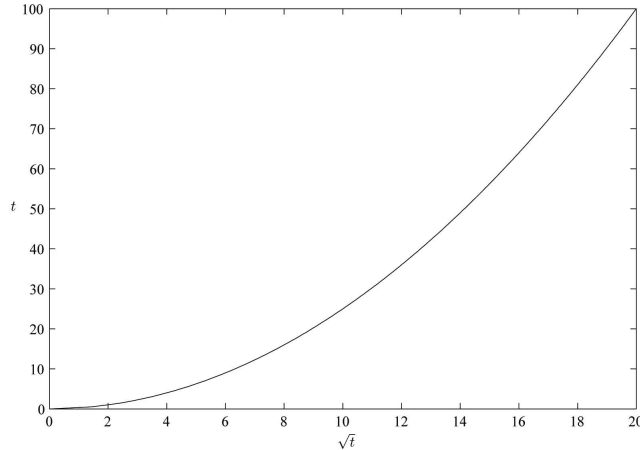


FIG. 1: The function representing the Fermat real $dt_2 \in D_3$

This inversion of abscissa and ordinate in the $\text{graph}_\delta(x)$ permits to represent this graph as a line tangent to the classical straight line \mathbb{R} and hence to have a better graphical picture. Finally, note that if $x \in \mathbb{R}$ is a standard real, then $N = 0$ and the $\text{graph}_\delta(x)$ is a vertical line passing through ${}^\circ x = x$.

The following theorem permits to represent geometrically the Fermat reals

Theorem 35. If $\delta \in \mathbb{R}_{>0}$, then the function

$$x \in \bullet\mathbb{R} \mapsto \text{graph}_\delta(x) \subset \mathbb{R}^2$$

is injective. Moreover if $x, y \in \bullet\mathbb{R}$, then we can find $\delta \in \mathbb{R}_{>0}$ (depending on x and y) such that

$$x < y$$

if and only if

$$\forall p, q, t : (p, t) \in \text{graph}_\delta(x) , (q, t) \in \text{graph}_\delta(y) \implies p < q \quad (33)$$

Proof: The application $\rho(x) := \text{graph}_\delta(x)$ for $x \in \bullet\mathbb{R}$ is well defined because it depends on the terms ${}^\circ x$, ${}^\circ x_i$ and $\omega_i(x)$ of the decomposition of x (see Theorem 5 and Definition 8).

Now, suppose that $\text{graph}_\delta(x) = \text{graph}_\delta(y)$, then

$$\forall t \in [0, \delta) : {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} = {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)}. \quad (34)$$

Let us consider the Fermat reals generated by these functions, i.e.

$$\begin{aligned} x' &:= \left[t \geq 0 \mapsto {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} \right]_{= \text{in } \bullet\mathbb{R}} \\ y' &:= \left[t \geq 0 \mapsto {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)} \right]_{= \text{in } \bullet\mathbb{R}} \end{aligned}$$

then the decompositions of x' and y' are exactly the decompositions of x and y

$$x' = {}^\circ x + \sum_{i=1}^N {}^\circ x_i dt_{\omega_i(x)} = x \quad (35)$$

$$y' = {}^\circ y + \sum_{j=1}^M {}^\circ y_j dt_{\omega_j(y)} = y. \quad (36)$$

But from (34) it follows $x' = y'$ in $\bullet\mathbb{R}$, and hence also $x = y$ from (35) and (36).

Now suppose that $x < y$, then, using the same notations of the previous part of this proof, we have also $x' = x$ and $y' = y$ and hence

$$x' = {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} < {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)} = y'.$$

We apply Theorem 28 obtaining that locally $x'_t < y'_t$, i.e.

$$\exists \delta > 0 : \forall {}^0t \geq 0 : {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} < {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)}.$$

This is an equivalent formulation of (33), and, because of Theorem 28 it is equivalent to $x' = x < y' = y$.

■

Example. In figure 2 we have the representation of some first order infinitesimals.

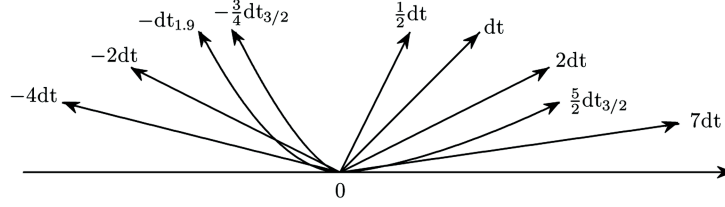


FIG. 2: Some first order infinitesimals

The arrows are justified by the fact that the representing function (32) is defined on $\mathbb{R}_{\geq 0}$ and hence has a clear first point and a direction. The smaller is $\alpha \in (0, 1)$ and the nearer is the representation of the product αdt , to the vertical line passing through zero, which is the representation of the standard real $x = 0$. Finally, recall that $dt_k \in D$ if and only if $1 \leq k < 2$.

If we multiply two infinitesimals we obtain a smaller number, hence one whose representation is nearer to the vertical line passing through zero, as represented in figure 3

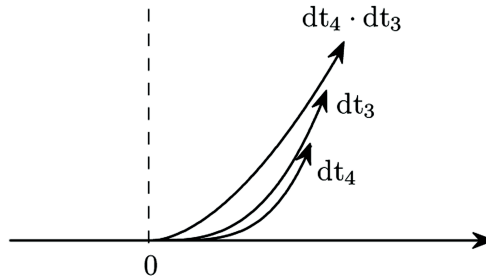


FIG. 3: The product of two infinitesimals

In figure 4 we have a representation of some infinitesimals of order greater than 1. We can see that the greater is the infinitesimal $h \in D_a$ (with respect to the order relation \leq defined in $\bullet\mathbb{R}$) and the higher is the order of intersection of the corresponding line $\text{graph}_\delta(h)$.

Finally, in figure 5 we represent the order relation on the basis of Theorem 35.

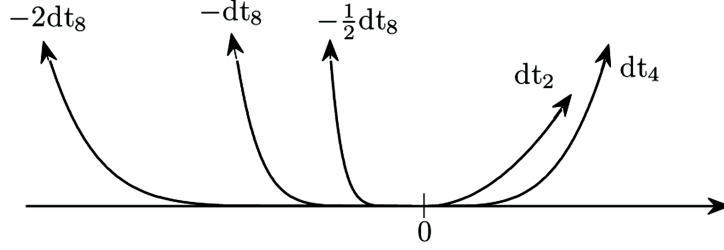
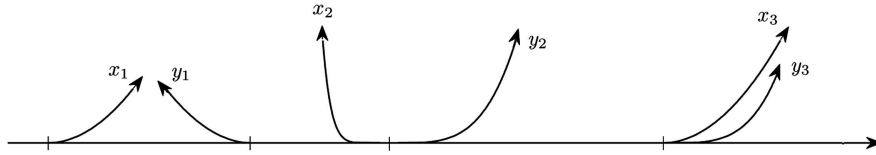


FIG. 4: Some higher order infinitesimals

FIG. 5: Different cases in which $x_i < y_i$

Intuitively, the method to see if $x < y$ is to look at a suitably small neighborhood (i.e. at a suitably small $\delta > 0$) at $t = 0$ of their representing lines $\text{graph}_\delta(x)$ and $\text{graph}_\delta(y)$: if, with respect to the horizontal directed straight line, the curve $\text{graph}_\delta(x)$ comes before the curve $\text{graph}_\delta(y)$, then x is less than y .

XIV. SOME ELEMENTARY EXAMPLES

The elementary examples presented in this section want to show, in a few rows, the simplicity of the algebraic calculus of nilpotent infinitesimals. Here “simplicity” means that the dialectic with the corresponding informal calculations, used e.g. in engineering or in physics, is really faithful. The importance of this dialectic can be glimpsed both as a proof of the flexibility of the new language, but also for researches in artificial intelligence like automatic differentiation theories (see e.g. Griewank [18] and references therein). Last but not least, it may also be important for didactic or historical researches. Several examples are directly taken from analogous of Bell [3] and the reader is strongly invited to compare the two theories in these cases. In particular, in our point of view, is not positive, like in some parts of Bell [3], to return back to a non rigorous use of infinitesimals. Mathematical theories of infinitesimals, like our ring of Fermat reals or NSA or SIA, are great opportunities to avoid several fallacies of the informal approach (our discussion in Section X is a clear example), and

to advance further, with the new knowledge originating from the rigorous theory, opening the possibility to use infinitesimal methods in more general, and less intuitive, frameworks (like e.g. infinite dimensional spaces of mappings, see Giordano [16]). Once again, the key point is the dialectic between formal and informal thoughts and not a single part only.

A. The heat equation.

In this and the following section we simply use the language of $\bullet\mathbb{R}$ to reformulate the corresponding deductions of Vladimirov [28]. Let us consider a body (identified with its localization) $B \subseteq \mathbb{R}^3$ and denote with $I_B := \text{int}(B)$ its interior. On I_B are given the smooth functions $\rho : I_B \rightarrow \mathbb{R}$, $c : I_B \rightarrow \mathbb{R}$ and $k : I_B \rightarrow \mathbb{R}$, interpreted respectively as the mass density, the specific heat capacity and the coefficient of thermal conductivity. Let us note that assuming these functions as defined on I_B without any favored direction corresponds physically to assume that B is an isotropous body. Moreover, let $u : I_B \times [0, +\infty) \rightarrow \mathbb{R}$ be the smooth function representing the temperature of the body B at each point $x \in I_B$ and time $t \in [0, +\infty)$. To deduce the heat diffusion equation we fix an internal point $x \in I_B$ and an infinitesimal volume V . More precisely, we say that a subset of $\bullet\mathbb{R}^3$ of the form

$$V = V(x, \delta\vec{x}) = \{y \in \bullet\mathbb{R}^3 \mid -\delta x_i \leq 2(y - x) \cdot \vec{e}_i \leq \delta x_i \quad \forall i = 1, 2, 3\} \quad (37)$$

is an infinitesimal parallelepiped if $\delta v := \delta x_1 \cdot \delta x_2 \cdot \delta x_3 \in D_\infty$, i.e. if the corresponding volume is an infinitesimal of some order. Here $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is the natural base of \mathbb{R}^3 and notations of the form $\delta y \in \bullet\mathbb{R}$ are only useful to underline that the infinitesimal increment is associated to the variable y : here δ is not an operator and we use it instead of the common dy to avoid confusion with our dy introduced in Definition 8. Because $x \in I_B$, the inclusion $V \subseteq \bullet B$ follows, so that V can be thought as the sub-body of B corresponding to the infinitesimal parallelepiped parallel to coordinate axis and centered at x . This sub-body V interacts thermally with its complement $\mathcal{C}V := \bullet B \setminus V$ and with external sources of heat. In the infinitesimal time interval $\delta t \in D_\infty$, the sub-body V exchanges with its complement $\mathcal{C}V$ the heat flowing perpendicularly to the surface of V (Fourier's law):

$$Q_{\mathcal{C}V, V} = \delta t \cdot \sum_{i=1}^3 \delta s_i \cdot \left[k(x + \delta\vec{h}_i) \cdot \frac{\partial u}{\partial \vec{e}_i}(x + \delta\vec{h}_i, t) - k(x - \delta\vec{h}_i) \cdot \frac{\partial u}{\partial \vec{e}_i}(x - \delta\vec{h}_i, t) \right], \quad (38)$$

where $\delta \vec{h}_i := \frac{1}{2} \delta x_i \cdot \vec{e}_i \in \bullet \mathbb{R}^3$ and $\delta s_i := \prod_{j \neq i} \delta x_j \in \bullet \mathbb{R}$. Choosing the infinitesimals so that

$$\delta v \cdot \delta t \in D,$$

we have that $\delta t \cdot \delta s_i \cdot (\delta x_i)^2 = \delta t \cdot \delta v \cdot \delta x_i = 0$ from Theorem 12 (e.g. we can choose $\delta x_i = dt_6$ and $\delta t = dt_2$). From this and the use of infinitesimal Taylor's formula in (38), simple calculations give

$$Q_{CV,V} = \operatorname{div} [k \cdot \operatorname{grad}(u)](x, t) \cdot \delta v \cdot \delta t. \quad (39)$$

Of course, this calculations correspond to the infinitesimal version of the Gauss-Ostrogradskij theorem. Interacting thermally with external sources, the sub-body V exchanges the heat

$$Q_{\text{ext},V} = F(x, t) \cdot \delta v \cdot \delta t, \quad (40)$$

where $F : I_B \times [0, +\infty) \longrightarrow \mathbb{R}$ is a smooth function representing the intensity of the thermal sources. The total heat $Q_{CV,V} + Q_{\text{ext},V}$ corresponds to an increasing of temperature of V equal to $u(x, t + \delta t) - u(x, t)$ and hence to an exchange of heat with the environment equal to

$$Q_{\text{env},V} = [u(x, t + \delta t) - u(x, t)] \cdot c(x) \cdot \rho(x) \cdot \delta v = Q_{CV,V} + Q_{\text{ext},V}. \quad (41)$$

From this and (39), (40), the infinitesimal Taylor's formula and the cancellation law we obtain the conclusion:

$$c(x) \cdot \rho(x) \cdot \frac{\partial u}{\partial t}(x, t) = \operatorname{div} [k \cdot \operatorname{grad}(u)](x, t) + F(x, t).$$

To stress that the previous deduction is now completely rigorous we can now state the following theorem, without any mention to the physical interpretation:

Theorem 36. *Let $B \subseteq \mathbb{R}^d$ and $I_B := \operatorname{int}(B)$ its interior. Let us consider the smooth functions $\rho : I_B \longrightarrow \mathbb{R}$, $c : I_B \longrightarrow \mathbb{R}$, $k : I_B \longrightarrow \mathbb{R}$, $u : I_B \times [0, +\infty) \longrightarrow \mathbb{R}$ and $F : I_B \times [0, +\infty) \longrightarrow \mathbb{R}$. Finally let us consider a point $(x, t) \in I_B \times [0, +\infty)$ and define V , $Q_{CV,V}$, $Q_{\text{ext},V}$, $Q_{\text{env},V}$ as in (37), (38), (40) and (41), where $\delta v \cdot \delta t \in D$. Then it results*

$$Q_{\text{env},V} = Q_{CV,V} + Q_{\text{ext},V}$$

if and only if the following relation holds

$$c(x) \cdot \rho(x) \cdot \frac{\partial u}{\partial t}(x, t) = \operatorname{div} [k \cdot \operatorname{grad}(u)](x, t) + F(x, t).$$

Unfortunately, this statement does not sufficiently underline the great difference that takes place between the physical content in the definition of $Q_{CV,V}$, i.e. the Fourier's law, and that in the definition of $Q_{\text{ext},V}$. In an axiomatic framework for thermodynamics (see e.g. Truesdell [27]), the notion of heat flux Q_{AB} going from a body A to a body B can be taken as primitive; in that case (38) becomes an important assumption, whereas (40) is simply the definition of the intensity $F(x, t) = \frac{Q_{\text{ext},V}}{\delta v \cdot \delta t}$.

B. Electric dipole.

In elementary Physics, an electric dipole is usually defined as “*a pair of charges with opposite sign placed at a distance d very less than the distance r from the observer*”. Conditions like $r \gg d$ are frequently used in Physics and very often we obtain a correct formalization if we ask $d \in \bullet\mathbb{R}$ infinitesimal but $r \in \mathbb{R} \setminus \{0\}$, i.e. r finite. Thus we can define an electric dipole as a pair (p_1, p_2) of electric particles, with charges of equal intensity but with opposite sign such that their mutual distance at every time t is a first order infinitesimal:

$$\forall t : |p_1(t) - p_2(t)| =: |\vec{d}_t| =: d_t \in D. \quad (42)$$

In this way we can calculate the potential at the point x using the properties of D and using the hypothesis that r is finite and not zero. In fact we have

$$\varphi(x) = \frac{q}{4\pi\epsilon_0} \cdot \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \quad \vec{r}_i := x - p_i$$

and if $\vec{r} := \vec{r}_2 - \frac{\vec{d}}{2}$ then

$$\frac{1}{r_2} = \left(r^2 + \frac{d^2}{4} + \vec{r} \cdot \vec{d} \right)^{-1/2} = r^{-1} \cdot \left(1 + \frac{\vec{r} \cdot \vec{d}}{r^2} \right)^{-1/2}$$

because for (42) $d^2 = 0$. For our hypotheses on d and r we have that $\frac{\vec{r} \cdot \vec{d}}{r^2} \in D$ hence from the derivation formula

$$\left(1 + \frac{\vec{r} \cdot \vec{d}}{r^2} \right)^{-1/2} = 1 - \frac{\vec{r} \cdot \vec{d}}{2r^2}$$

In the same way we can proceed for $1/r_1$, hence:

$$\begin{aligned} \varphi(x) &= \frac{q}{4\pi\epsilon_0} \cdot \frac{1}{r} \cdot \left(1 + \frac{\vec{r} \cdot \vec{d}}{2r^2} - 1 + \frac{\vec{r} \cdot \vec{d}}{2r^2} \right) = \\ &= \frac{q}{4\pi\epsilon_0} \cdot \frac{\vec{r} \cdot \vec{d}}{r^3} \end{aligned}$$

The property $d^2 = 0$ is also used in the calculus of the electric field and for the moment of momentum.

C. Newtonian limit in Relativity.

Another example in which we can formalize a condition like $r \gg d$ using the previous ideas is the Newtonian limit in Relativity; in it we can suppose to have

- $\forall t : v_t \in D_2 \quad \text{and} \quad c \in \mathbb{R}$
- $\forall x \in M_4 : g_{ij}(x) = \eta_{ij} + h_{ij}(x) \quad \text{with} \quad h_{ij}(x) \in D.$

where $(\eta_{ij})_{ij}$ is the matrix of the Minkowski's metric. This conditions can be interpreted as $v_t \ll c$ and $h_{ij}(x) \ll 1$ (low speed with respect to the speed of light and weak gravitational field). In this way we have, e.g. the equalities:

$$\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1 + \frac{v^2}{2c^2} \quad \text{and} \quad \sqrt{1 - h_{44}(x)} = 1 - \frac{1}{2} h_{44}(x).$$

D. Linear differential equations.

Let

$$L(y) := A_0 \frac{d^N y}{dt^N} + \dots + A_{N-1} \frac{dy}{dt} + A_N \cdot y = 0$$

be a linear differential equation with constant coefficients. Once again we want *to discover* independent solutions in case the characteristic polynomial has multiple roots e.g.

$$(r - r_1)^2 \cdot (r - r_3) \cdot \dots \cdot (r - r_N) = 0.$$

The idea is that in $\bullet\mathbb{R}$ we have $(r - r_1)^2 = 0$ also if $r = r_1 + h$ with $h \in D$. Thus $y(t) = e^{(r_1+h)t}$ is a solution too. But $e^{(r_1+h)t} = e^{r_1 t} + ht \cdot e^{r_1 t}$, hence

$$\begin{aligned} L[e^{(r_1+h)t}] &= 0 \\ &= L[e^{r_1 t} + ht \cdot e^{r_1 t}] \\ &= L[e^{r_1 t}] + h \cdot L[t \cdot e^{r_1 t}] \end{aligned}$$

We obtain $L[t \cdot e^{r_1 t}] = 0$, that is $y_1(t) = t \cdot e^{r_1 t}$ must be a solution. Using k -th order infinitesimals we can deal with other multiple roots in a similar way.

E. Circle of curvature.

A simple application of the infinitesimal Taylor's formula is the parametric equation for the circle of curvature, that is the circle with second order osculation with a curve $\gamma : [0, 1] \longrightarrow \mathbb{R}^3$. In fact if $r \in (0, 1)$ and $\dot{\gamma}_r$ is a unit vector, from the second order infinitesimal Taylor's formula we have

$$\forall h \in D_2 : \gamma(r+h) = \gamma_r + h \dot{\gamma}_r + \frac{h^2}{2} \ddot{\gamma}_r = \gamma_r + h \vec{t}_r + \frac{h^2}{2} c_r \vec{n}_r \quad (43)$$

where \vec{n} is the unit normal vector, \vec{t} is the tangent one and c_r the curvature. But once again from Taylor's formula we have $\sin(ch) = ch$ and $\cos(ch) = 1 - \frac{c^2 h^2}{2}$. Now it suffices to substitute h and $\frac{h^2}{2}$ from these formulas into (43) to obtain the conclusion

$$\forall h \in D_2 : \gamma(r+h) = \left(\gamma_r + \frac{\vec{n}_r}{c_r} \right) + \frac{1}{c_r} \cdot [\sin(c_r h) \vec{t}_r - \cos(c_r h) \vec{n}_r].$$

In a similar way we can prove that any $f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ can be written $\forall h \in D_k$ as

$$f(h) = \sum_{n=0}^k a_n \cdot \cos(nh) + \sum_{n=0}^k b_n \cdot \sin(nh),$$

so that now the idea of the Fourier series comes out in a natural way.

F. Commutation of differentiation and integration.

This example derives from Kock [20], Lavendhomme [22]. Suppose we want *to discover* the derivative of the function

$$g(x) := \int_{\alpha(x)}^{\beta(x)} f(x, t) dt \quad \forall x \in \mathbb{R}$$

where α , β and f are smooth functions. We can see g as a composition of smooth functions, hence we can apply the derivation formula, i.e. Theorem 21:

$$\begin{aligned} g(x+h) &= \int_{\alpha(x+h)}^{\beta(x+h)} f(x+h, t) dt = \\ &= \int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} f(x, t) dt + h \cdot \int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} \frac{\partial f}{\partial x}(x, t) dt + \\ &+ \int_{\alpha(x)}^{\beta(x)} f(x, t) dt + h \cdot \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x, t) dt + \\ &+ \int_{\beta(x)}^{\beta(x)+h\beta'(x)} f(x, t) dt + h \cdot \int_{\beta(x)}^{\beta(x)+h\beta'(x)} \frac{\partial f}{\partial x}(x, t) dt. \end{aligned}$$

Now we use $h^2 = 0$ to obtain e.g. (see Corollary 23):

$$h \cdot \int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} \frac{\partial f}{\partial x}(x, t) dt = -h^2 \cdot \alpha'(x) \cdot \frac{\partial f}{\partial x}(\alpha(x), t) = 0$$

and

$$\int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} f(x, t) dt = -h \cdot \alpha'(x) \cdot f(\alpha(x), t).$$

Calculating in an analogous way similar terms we finally obtain the well known conclusion. Note that the final formula comes out by itself so that we have “*discovered*” it and not simply we have proved it. From the point of view of artificial intelligence or from the didactic point of view, surely this discovering is not a trivial result.

G. Schwarz’s theorem.

Using nilpotent infinitesimals we can obtain a simple and meaningful proof of Schwarz’s theorem. This simple example aims to show how to manage some differences between our setting and SDG. Let $f : V \longrightarrow E$ be a \mathcal{C}^2 function between spaces of type $V = \mathbb{R}^m$, $E = \mathbb{R}^n$ and $a \in V$, we want to prove that $d^2f(a) : V \times V \longrightarrow E$ is symmetric. Take

$$k \in D_2$$

$$h, j \in \mathcal{D}_\infty$$

$$jkh \in D_{\neq 0}$$

(e.g. we can take $k_t = dt_2, h_t = j_t = dt_4$ so that $jkh = dt$, see also Theorem 12). Using $k \in D_2$, we have

$$\begin{aligned} j \cdot f(x + hu + kv) &= \\ &= j \cdot \left[f(x + hu) + k \partial_v f(x + hu) + \frac{k^2}{2} \partial_v^2 f(x + hu) \right] \\ &= j \cdot f(x + hu) + jk \cdot \partial_v f(x + hu) \end{aligned} \tag{44}$$

where we used the fact that $k^2 \in D$ and j infinitesimal imply $jk^2 = 0$. Now we consider that $jkh \in D$ so that any product of type $jkhi$ is zero for every $i \in D_\infty$, so we obtain

$$jk \cdot \partial_v f(x + hu) = jk \cdot \partial_v f(x) + jkh \cdot \partial_u(\partial_v f)(x). \tag{45}$$

But $k \in D_2$ and $jk^2 = 0$ hence

$$j \cdot f(x + kv) - j \cdot f(x) = jk \cdot \partial_v f(x).$$

Substituting this in (45) and hence in (44) we obtain

$$\begin{aligned} j \cdot [f(x + hu + kv) - f(x + hu) - f(x + kv) + f(x)] = \\ = jkh \cdot \partial_u(\partial_v f)(x). \end{aligned} \quad (46)$$

The left hand side of this equality is symmetric in u, v , hence changing them we have

$$jkh \cdot \partial_u(\partial_v f)(x) = jkh \cdot \partial_v(\partial_u f)(x)$$

and thus we obtain the conclusion because $jkh \neq 0$ and $\partial_u(\partial_v f)(x), \partial_v(\partial_u f)(x) \in E$. From (46) it follows directly the classical limit relation

$$\lim_{t \rightarrow 0^+} \frac{f(x + h_t u + k_t v) - f(x + h_t u) - f(x + k_t v) + f(x)}{h_t k_t} = \partial_u \partial_v f(x)$$

H. Area of the circle and volumes of revolution.

A more or less meaningful proof of the familiar formula for the area of a circle depends on what axioms are assumed and how much general the definitions are. In this example we want to show the possibility to define suitable smooth functions using an infinitesimal property. Let us assume the axioms for the real field \mathbb{R} ; prove from them the existence of the smooth functions \sin and \cos ; define π as a suitable zero of these functions (see e.g. Prodi [24], Silov [26]) and define the length of an arc of circle of radius r , parametrized by $x(\theta) = r \cdot \cos(\theta)$ and $y(\theta) = r \cdot \sin(\theta)$, as the unique function s that verifies

$$[s(\theta + k) - s(\theta)]^2 = [x(\theta + k) - x(\theta)]^2 + [y(\theta + k) - y(\theta)]^2 \quad \forall \theta \in \mathbb{R} \quad \forall k \in D_2 \quad (47)$$

$$s(0) = 0. \quad (48)$$

This definition can be justified in the usual way using a (second order!) infinitesimal right-angled triangle. The uniqueness of s follows from (47) and (48), the smoothness of x and y , the second order infinitesimal Taylor's formula and the cancellation law (Theorem 17):

$$k^2 \cdot \dot{s}(\theta) = \dot{x}(\theta) \cdot k^2 + \dot{y}(\theta) \cdot k^2 \quad \forall k \in D_2.$$

From this and (48) we obtain the usual formula for s that, in our particular case, gives $s(\theta) = r \cdot \theta$. Now we can think the area $A(\theta + h) - A(\theta)$ of a first order infinitesimal sector of the circle as the area of the isosceles triangle with sides of length r and base $s(\theta + h) - s(\theta)$. In fact, if $P(\theta) = (r \sin \theta, r \cos \theta)$, then $P(\theta + h) = P(\theta) + h \cdot \vec{t}(\theta)$, where \vec{t} is the tangent

vector, so that in $[\theta, \theta + h]$, $h \in D$, the circle is made of linear segments. Therefore, the area $A(\theta)$ can be defined as the unique function that verifies

$$A(\theta + h) - A(\theta) = \frac{1}{2} [s(\theta + h) - s(\theta)] \cdot r \cos \left(\frac{h}{2} \right) \quad \forall \theta \in \mathbb{R} \quad \forall h \in D$$

$$A(0) = 0.$$

From this and the derivation formula we get

$$h \cdot A'(\theta) = \frac{1}{2} h r \cdot s'(\theta)$$

$$A(\theta) = \frac{1}{2} \int_0^\theta r \cdot s(u) \, du.$$

In our case we get $A(\theta) = \frac{1}{2} r^2 \cdot \theta$ and hence the searched formula for $\theta = 2\pi$.

Analogously we can prove the familiar formula for volumes of revolution of a parametrized curve $\gamma(u) = (x(u), y(u))$, $u \in [a, b]$, around the x -axis. Let us define the volume as the unique smooth function V that verifies

$$V(u + h) - V(u) = h \cdot \pi \cdot y(u)^2 + \frac{1}{2} [h \cdot \pi \cdot y(u + h)^2 - h \cdot \pi \cdot y(u)^2] \quad (49)$$

$$V(0) = 0 \quad (50)$$

for every $u \in [a, b]$ and $h \in D$. This definition can be intuitively justified saying that the volume of the sector of revolution between u and $u + h$ can be calculated as the sum of the cylinder of radius $y(u)$ and height h plus one half of the difference between the cylinder of radius $y(u + h)$ and height h and that of radius $y(u)$ and the same height. Implicitly, we are using the straightness of the curve γ in $[u, u + h]$. From (49) and the property $h^2 = 0$ we easily obtain that $V'(u) = \pi \cdot y(u)^2$ and hence the usual formula using (50).

I. Curvature.

Let us consider the usual smooth parametrized curve $\gamma(u) = (x(u), y(u))$ for $u \in [a, b]$. Let $\varphi(u) \in [0, \pi]$ be the non-oriented angle (i.e. the one defined by the scalar product) between the tangent vector $\vec{t} = (\dot{x}, \dot{y})$ and the unit vector \vec{i} of the x -axis, so that

$$\sqrt{\dot{x}^2 + \dot{y}^2} \cdot \cos \varphi = \dot{x}.$$

Multiplying this equality by $\sin \varphi$ we easily obtain

$$\dot{y} \cdot \cos \varphi = \dot{x} \cdot \sin \varphi. \quad (51)$$

It is well known that the curvature of γ at the point $u \in [a, b]$ can be calculated as the rate of change of the non-oriented angle $\varphi(u)$ with respect to an infinitesimal variation in arc length $s(u)$ defined by the analogous of (47) and (48). These “rate of changes” can be defined in $\bullet\mathbb{R}$ as the unique (if it exists) standard $c(u) \in \mathbb{R}$ defined by

$$c(u) \cdot [s(u+h) - s(u)] = \varphi(u+h) - \varphi(u) \quad \forall h \in D.$$

Indeed, from the cancellation law, i.e. Theorem 17, there exists at most one such $c(u) \in \mathbb{R}$ verifying this property. Because of this uniqueness we can also use the notation

$$c(u) = \frac{\varphi(u+h) - \varphi(u)}{s(u+h) - s(u)}. \quad (52)$$

These ratios generalize the usual ratios between real numbers (see Giordano [16] for more details). From (52) and the derivation formula we get $c(u) = \frac{h \cdot \varphi'(u)}{h \cdot s'(u)} = \frac{\varphi'(u)}{s'(u)}$ whatever $h \in D_{\neq 0}$ we choose. From this and the relation (51) (without using infinitesimals, but using standard differential calculus) we can obtain the usual formula $c = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$ at each point $u \in [a, b]$ where $\varphi(u) \neq \frac{\pi}{2}$ and $\dot{\gamma}(u) \neq \underline{0}$.

J. Stretching of a spring (and center of pressure).

If $f : [a, b] \longrightarrow \mathbb{R}$ is a smooth function and we define $J(x) := \int_0^x f(s) ds$, then Corollary 23 and a trivial calculation with the derivation formula give

$$J(x+h) - J(x) = \frac{1}{2} [f(x+h) + f(x)] \quad \forall h \in D. \quad (53)$$

The right-hand side of (53) is interpreted as the average value of f in the infinitesimal interval $[x, x+h]$. Analogous equalities can be obtain in the d -dimensional case using suitable generalizations of the above cited corollary: e.g. if $d = 2$ we have to use

$$\int_0^h \int_0^k f(x, y) dx dy = hk \cdot f(0, 0) \quad \forall h, k \in D_\infty : h \cdot k \in D.$$

These equalities are used by Bell [3] to calculate the center of pressure of a plane area and the work done in stretching a spring. The meaningfulness of such examples is however doubtful because they can be summarized saying: assume to have a smooth J satisfying (53); deduce from this and from the assumption $J(0) = 0$ that $J'(x) = f(x)$. There is no real use of infinitesimals in this type of reasoning in every case where the definition $J(x) := \int_0^x f(s) ds$ is customary, like in the cited examples.

K. The wave equation.

The deduction of the wave equation in the framework of Fermat reals is very interesting for two main reasons. Firstly, in the classical deduction (see e.g. Vladimirov [28]) there are some approximations tied with Hook's law. Is it possible to make them rigorous using $\bullet\mathbb{R}$? Do we gain something using this increased rigour? E.g.: how can we formalize the approximated equalities used in the classical deduction? In what a sense is the wave equation an approximated equality valid for small oscillations only?

Secondly, at the end of our deduction we will stress the physical principles as important mathematical assumptions of a suitable theorem. We are hence naturally taken to ask if these natural assumptions (some of which formulated using the infinitesimals of $\bullet\mathbb{R}$) really have a model. In this way, we will see that no standard smooth function can satisfy these hypothesis, but we are forced to consider a non-standard one. E.g. $f(x) = h \cdot \sin(x)$ for $x \in \bullet\mathbb{R}$ and $h \in D_\infty$ is an example of a non-standard smooth function; let us note that it is obtained by the standard smooth function $g(y, x) := y \cdot \sin(x)$, $x, y \in \mathbb{R}$, by extension to $\bullet\mathbb{R}^2$ and fixing one of its variables to a non-standard parameter $h \in D_\infty$:

$$f(x) = \bullet g(h, x) \quad \forall x \in \bullet\mathbb{R}.$$

This will motivate strongly the further development of the theory of Fermat reals, in the direction of a more general theory including also these new smooth non-standard functions.

Let us start considering a string making small transversal oscillations around its equilibrium position located on the interval $[a, b]$ of the x axis, for $a, b \in \mathbb{R}$, $a < b$. By hypotheses, string's position $s_t \subseteq \bullet\mathbb{R}^2$ is always represented by the graph of a given curve $\gamma : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}^2$ (where $[a, b] = \{x \in \bullet\mathbb{R} \mid a \leq x \leq b\}$ and $[0, +\infty) = \{x \in \bullet\mathbb{R} \mid 0 \leq x\}$; in the following, we will always use these notations for intervals to identify the corresponding subsets of $\bullet\mathbb{R}$, and not of \mathbb{R} , and we will also use the notation $\gamma_{xt} := \gamma(x, t)$):

$$s_t = \{\gamma_{xt} \in \bullet\mathbb{R}^2 \mid a \leq x \leq b\} \quad \forall t \in [0, +\infty).$$

Moreover, the curve γ is supposed to be injective with respect to the parameter $x \in (a, b)$:

$$\gamma_{x_1 t} \neq \gamma_{x_2 t} \quad \forall t \in [0, +\infty) \quad \forall x_1, x_2 \in (a, b) : x_1 \neq x_2,$$

so that the order relation on (a, b) implies an order relation on the support s_t . For every

pair of points $p = \gamma_{x_p t}$, $q = \gamma_{x_q t} \in s_t$ on the string at time t , we can define the sub-bodies:

$$\begin{aligned}\overrightarrow{p} &:= \{\gamma_{xt} \mid x_p \leq x \leq b\} \\ \overleftarrow{p} &:= \{\gamma_{xt} \mid a \leq x \leq x_p\} \\ \overrightarrow{pq} &:= \{\gamma_{xt} \mid x_p \leq x \leq x_q\}\end{aligned}$$

corresponding respectively to the parts of the string that follows the point $p \in s_t$, that precedes the same point and that lies between the point $p \in s_t$ and the point $q \in s_t$. It is usually implicitly clear that e.g. every sub-body of the form \overrightarrow{p} exerts a force on each sub-body with which it is in contact, i.e. of the form \overrightarrow{pq} or \overleftarrow{p} . Moreover, the force $\mathbf{F}(A, B) \in \bullet\mathbb{R}^2$ that the sub-body A exerts on the sub-body B verifies the following equalities (see e.g. Truesdell [27]):

$$\mathbf{F}(\overrightarrow{pq}, \overleftarrow{p}) = \mathbf{F}(\overrightarrow{p}, \overleftarrow{p}) \quad (54)$$

$$\mathbf{F}(\overrightarrow{q}, \overrightarrow{pq}) = \mathbf{F}(\overrightarrow{q}, \overleftarrow{q}) \quad (55)$$

$$\mathbf{F}(\overleftarrow{p}, \overrightarrow{pq}) = -\mathbf{F}(\overrightarrow{pq}, \overleftarrow{p}) \quad (\text{action-reaction principle}) \quad (56)$$

for every pair of points $p, q \in s_t$ and every time $t \in [0, +\infty)$. Using this formalism, the tension at the point $\gamma_{xt} \in s_t$ at time $t \in [0, +\infty)$ can now be defined in the following way

$$\mathbf{T}(x, t) := \mathbf{F}(\overrightarrow{\gamma_{xt}}, \overleftarrow{\gamma_{xt}}). \quad (57)$$

Now, let us consider the infinitesimal sub-body $\overrightarrow{x, x + \delta x} := \overrightarrow{\gamma_{xt} \gamma_{x+\delta x, t}} \subseteq s_t$ located at time t between the points $\gamma_{xt} \in s_t$ and $\gamma_{x+\delta x, t} \in s_t$, where $\delta x \in D$ is a generic first order infinitesimal. On this infinitesimal sub-body, mass forces of linear density $\mathbf{G} : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}^2$ act, so that Newton's law can be written as

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} = \mathbf{F}(\overleftarrow{\gamma_{xt}}, \overrightarrow{x, x + \delta x}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x, t}}, \overrightarrow{x, x + \delta x}) + \mathbf{G} \cdot \rho \cdot \delta x, \quad (58)$$

where $\rho : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}$ is the linear mass density and where, if not otherwise indicated, all the functions are calculated at $(x, t) \in (a, b) \times [0, +\infty)$. Of course, the contact forces appearing in Newton's law are due to the interaction of the infinitesimal sub-body with other sub-bodies in contact with its border

$$\partial \left[\overrightarrow{x, x + \delta x} \right] = \{\gamma_{xt}, \gamma_{x+\delta x, t}\} \subseteq \bullet\mathbb{R}^2.$$

Using action-reaction principle (56) and the equality (55), with $q = \gamma_{x+\delta x, t}$ and $p = \gamma_{xt}$ so that $\overrightarrow{pq} = \overrightarrow{x, x + \delta x}$, from (58) we have

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} = -\mathbf{F}(\overrightarrow{x, x + \delta x}, \overleftarrow{\gamma_{xt}}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x, t}}, \overleftarrow{\gamma_{x+\delta x, t}}) + \mathbf{G} \cdot \rho \cdot \delta x.$$

Using (54) and the definition (57) of tension we get

$$\begin{aligned} \rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} &= -\mathbf{F}(\overrightarrow{\gamma_{xt}}, \overleftarrow{\gamma_{xt}}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x, t}}, \overleftarrow{\gamma_{x+\delta x, t}}) + \mathbf{G} \cdot \rho \cdot \delta x \\ &= -\mathbf{T}(x, t) + \mathbf{T}(x + \delta x, t) + \mathbf{G} \cdot \rho \cdot \delta x. \end{aligned} \quad (59)$$

Up to this point of the deduction we have not used neither the hypotheses of small oscillations nor that of transversal oscillations. The second one can be easily introduced with the hypotheses

$$\mathbf{G}(x, t) \cdot \vec{e}_1 = 0 \quad \forall x, t, \quad (60)$$

where (\vec{e}_1, \vec{e}_2) are the axis unit vectors. Using the notation $\varphi(x, t)$ for the non-oriented angle between the tangent unit vector $\mathbf{t}(x, t)$ at the point γ_{xt} and the x axes (see (51)), the hypotheses of small oscillations can be formalized with the assumption

$$\varphi(x, t) \in D \quad \forall x, t. \quad (61)$$

This will permit to reproduce the classical deduction in the most faithful way (even if, as we will see later, a weaker assumption can be considered). Moreover, in the classical deduction of the wave equation, one considers only curves of the form $\gamma_{xt} = (x, u(x, t))$. In this way from (51) and the derivation formula we have

$$\begin{aligned} \frac{\partial \gamma_2}{\partial x} \cdot \cos \varphi &= \sin \varphi \\ \frac{\partial \gamma_2}{\partial x} &= \varphi \in D \end{aligned}$$

so that $\left(\frac{\partial \gamma_2}{\partial x}\right)^2 = 0$ and hence the total length of the string becomes:

$$L = \int_a^b \sqrt{1 + \left[\frac{\partial \gamma_2}{\partial x}(x, t)\right]^2} dx = b - a \quad \forall t \in [0, +\infty). \quad (62)$$

By Hook's law, this justifies that the tension can be assumed to have a constant modulus T , not depending neither by the position x nor by the time t :

$$\mathbf{T}(x, t) = T \cdot \mathbf{t}(x, t) \quad \forall x \in (a, b) \quad \forall t \in [0, +\infty). \quad (63)$$

A tension \mathbf{T} parallel to the tangent vector is the second part of the hypothesis about non transversal oscillations of the string. Let us note explicitly that the only standard continuous function verifying the equality $L = b - a$ is the constant one, so the function $u : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}$ has to be understood as a non-standard one; later we will do further considerations about this important point. Projecting the equation (59) on the y axis, we obtain

$$\begin{aligned} \rho \cdot \delta x \cdot \frac{\partial^2 u}{\partial t^2} &= -T \cdot \mathbf{t}(x, t) \cdot \vec{e}_2 + T \cdot \mathbf{t}(x + \delta x, t) \cdot \vec{e}_2 + \mathbf{G} \cdot \vec{e}_2 \cdot \rho \cdot \delta x \\ &= -T \sin \varphi(x, t) + T \cdot \sin \varphi(x + \delta x, t) + G \cdot \rho \cdot \delta x. \end{aligned}$$

But $\sin \varphi = \varphi = \frac{\partial u}{\partial x}$ because $\varphi \in D$ is a first order infinitesimal, hence

$$\begin{aligned} \rho \cdot \delta x \cdot \frac{\partial^2 u}{\partial t^2} &= T \cdot \left[\frac{\partial u}{\partial x}(x + \delta x, t) - \frac{\partial u}{\partial x}(x, t) \right] + G \cdot \rho \cdot \delta x \\ &= \left[T \cdot \frac{\partial^2 u}{\partial x^2}(x, t) + G \cdot \rho \right] \cdot \delta x. \end{aligned} \tag{64}$$

We cannot use the cancellation law with $\delta x \in D$ to obtain the final result, because, as we mentioned above, the function $u(x, t) \in \bullet\mathbb{R}$ can assume non standard values, so it is time to clarify some points. As mentioned above, there does not exist a standard smooth function verifying all the assumptions or the physical principles we have used. Of course, everything depends by how we formalize the classical informal deduction used in elementary physics: e.g. we have chosen to use an equality sign in (62) instead of an approximated equality; anyway we have to consider that if we use \simeq to write (62), then the problem becomes how to make more precise, physically, numerically or mathematically, this approximation; moreover, if we use an approximation sign in (62), then we consistently must use the same sign both in (63) and therefore in the final wave equation. Nevertheless, smooth non standard functions can verify all the hypothesis and physical principles we have considered: e.g. the function $u(x, t) := u_0 \sin(x + \omega \cdot t)$ is one of these if the maximum amplitude $u_0 \in D$ and if ρ is constant, $G = 0$ and $T = \omega^2 \rho$.

Definition 37. If $X \subseteq \bullet\mathbb{R}^x$ and $Y \subseteq \bullet\mathbb{R}^y$ then we say that

$$f : X \longrightarrow Y \text{ is (non standard) smooth}$$

iff f maps X in Y and for every $x_0 \in X$ we can write

$$f(x) = \bullet g\langle p, x \rangle \quad \forall x \in \bullet V \cap X \tag{65}$$

for some

$$\begin{aligned} V &\text{ open in } \mathbb{R}^x \text{ such that } x_0 \in \bullet V \\ p &\in \bullet U, \text{ where } U \text{ is open in } \mathbb{R}^p \\ g &\in \mathcal{C}^\infty(U \times V, \mathbb{R}^y), \end{aligned}$$

where $\langle -, - \rangle : ([x]_\sim, [y]_\sim) \in \bullet U \times \bullet V \mapsto [(x, y)]_\sim \in \bullet(U \times V)$ (see Definition 4 for the relation \sim).

In other words locally a smooth function $f : X \longrightarrow Y$ from $X \subseteq \bullet \mathbb{R}^x$ to $Y \subseteq \bullet \mathbb{R}^y$ is constructed in the following way:

1. start with an ordinary standard function $g \in \mathcal{C}^\infty(U \times V, \mathbb{R}^y)$, with U open in \mathbb{R}^p and V open in \mathbb{R}^x . The space \mathbb{R}^p has to be thought as a space of parameters for the function g ;
2. consider its Fermat extension obtaining $\bullet g : \bullet(U \times V) \longrightarrow \bullet \mathbb{R}^y$;
3. consider the composition $\bullet g \circ \langle -, - \rangle : \bullet U \times \bullet V \longrightarrow \bullet \mathbb{R}^y$, where $\langle -, - \rangle$ is the isomorphism $\bullet U \times \bullet V \simeq \bullet(U \times V)$ defined by $\langle [x]_\sim, [y]_\sim \rangle = [(x, y)]_\sim$; we will always use the identification $\bullet U \times \bullet V = \bullet(U \times V)$, so we will write simply $\bullet g(p, x)$ instead of $\bullet g\langle p, x \rangle$.
4. fix a parameter $p \in \bullet U$ as a first variable of the previous composition, i.e. consider $\bullet g\langle p, - \rangle : \bullet V \longrightarrow \bullet \mathbb{R}^y$. Locally, the map f is of this form: $f = \bullet g\langle p, - \rangle = \bullet g(p, -)$.

Because $p = {}^\circ p + h$, with $h \in D_\infty$, applying the infinitesimal Taylor's formula to variable p for the function $\bullet g(p, x)$ it is not hard to prove the following Theorem, that clarifies further the form of these non standard smooth functions, because it states that they can be seen locally as “infinitesimal polynomials with smooth coefficients”:

Theorem 38. *Let $X \subseteq \bullet \mathbb{R}^x$ and $f : X \longrightarrow \bullet \mathbb{R}^n$ a map. Then it results that*

$$f : X \longrightarrow \bullet \mathbb{R}^n \text{ is non standard smooth}$$

if and only if for every $x_0 \in X$ we can write

$$f(x) = \sum_{\substack{|q| \leq k \\ q \in \mathbb{N}^d}} a_q(x) \cdot p^q \quad \forall x \in \bullet V \cap X, \quad (66)$$

for suitable:

1. $d, k \in \mathbb{N}$
2. $p \in D_k^d$
3. V open subset of \mathbb{R}^x such that $x_0 \in \bullet V$
4. $(a_q)_{\substack{|q| \leq k \\ q \in \mathbb{N}^d}}$ family of $\mathcal{C}^\infty(V, \mathbb{R}^n)$.

In other words, every smooth function $f : X \longrightarrow \bullet \mathbb{R}^n$ can be constructed locally starting from some “infinitesimal parameters”

$$p_1, \dots, p_d \in D_k$$

and from ordinary smooth functions

$$a_q \in \mathcal{C}^\infty(V, \mathbb{R}^n)$$

and using polynomial operation only with p_1, \dots, p_d and with coefficients $a_q(-)$. Roughly speaking, we can say that they are “infinitesimal polynomials with smooth coefficients. The polynomials variables act as parameters only”.

As it is natural to expect, several notions of differential and integral calculus, including their infinitesimal versions, can be extended to this type of new smooth function (for more details, see the preprint Giordano [16]), and these results will be presented in future works. In this sense, this deduction of the wave equation motivates strongly the future development of the theory of Fermat reals.

On the other hand, we have to understand what type of cancellation law we can apply to (64). For this end, we have to define the notion of equality up to k -th order infinitesimals:

Definition 39. Let $m = \circ m + \sum_{i=1}^N \circ m_i \cdot dt_{\omega_i(m)}$ be the decomposition of $m \in \bullet \mathbb{R}$ and $k \in \mathbb{R}_{\geq 0} \cup \{\infty\}$, then

$$\iota_k m := \iota_k(m) := \circ m + \sum_{\substack{i=1 \\ \omega_i(m) > k}}^N \circ m_i \cdot dt_{\omega_i(m)}.$$

Finally if $x, y \in \bullet \mathbb{R}$, we will say $x =_k y$ iff $\iota_k x = \iota_k y$ in $\bullet \mathbb{R}$, and we will read it as x is equal to y up to k -th order infinitesimals.

In other words, as it is easy to prove, we have

$$x =_k y \iff {}^\circ x = {}^\circ y \text{ and } \omega(x - y) \leq k.$$

Therefore, if we denote with

$$I_k := \{x \in D_\infty \mid \omega(x) \leq k\},$$

the set of all the infinitesimal of order less than or equal to k (let us note that $I_k \subset D_k$), then we have that $x =_k y$ if and only if $x - y \in I_k$. Equality up to k -th order infinitesimal is of course an equivalence relation and preserves all the ring operations of ${}^\bullet\mathbb{R}$. More in general these equalities are preserved by smooth functions $f : {}^\bullet\mathbb{R} \longrightarrow {}^\bullet\mathbb{R}$:

$$x =_k y \implies f(x) =_k f(y).$$

Using this notion, it is not hard to prove the following cancellation law up to k -th order infinitesimals.

Theorem 40. *Let $m \in {}^\bullet\mathbb{R}$, $n \in \mathbb{N}_{>0}$, $j \in \mathbb{N}^n \setminus \{0\}$ and $\alpha \in \mathbb{R}_{>0}^n$. Moreover let us consider $k \in \mathbb{R}$ defined by*

$$\frac{1}{k} + \sum_{i=1}^n \frac{j_i}{\alpha_i + 1} = 1 \quad (67)$$

then

$$1. \forall h \in D_{\alpha_1} \times \cdots \times D_{\alpha_n} : h^j \cdot m = h^j \cdot \iota_k m$$

$$2. \text{ If } h^j \cdot m = 0 \text{ for every } h \in D_{\alpha_1} \times \cdots \times D_{\alpha_n}, \text{ then } m =_k 0$$

E.g. if $n = 1$ and $\alpha_1 = j_1 = 1$ we have $k = 2$ and hence

$$\forall h \in D : h \cdot m = h \cdot \iota_2 m$$

$$(\forall h \in D : h \cdot m = 0) \iff m =_2 0. \quad (68)$$

Using (68) in (64) we obtain the final conclusion

$$\rho \cdot \frac{\partial^2 u}{\partial t^2} =_2 T \cdot \frac{\partial^2 u}{\partial x^2} + G \cdot \rho \quad \forall x \in (a, b) \quad \forall t \in (0, +\infty). \quad (69)$$

It is also interesting to note that not only small oscillations of the string implies (69), but the converse is also true: the equation (69) implies that necessary we must have small oscillations of the string, i.e. that $\varphi(x, t) \in D_\infty$. Moreover, using the equality $=_2$ up to second order infinitesimals, all the classical approximation tied with Hook's law, now become more clear. Indeed, we have the following

Theorem 41. Let $a, b \in \mathbb{R}$, with $a < b$; let $\gamma : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}^2$, $\rho : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}$ and $\mathbf{G}, \mathbf{T} : [a, b] \times [0, +\infty) \longrightarrow \bullet\mathbb{R}^2$ be non standard smooth functions and $T \in \bullet\mathbb{R}$ be an invertible Fermat real. Let us suppose that the first component γ_1 of the curve is of the form

$$\gamma_1(x, t) = [1 + \alpha(t)] \cdot x + \beta(t) \quad \forall x, t, \quad (70)$$

with $\alpha(t) \in I_2$. Then the unit tangent vector $\mathbf{t}(x, t)$ to the curve γ exists and we can further suppose that the relations

$$\mathbf{T}(x, t) =_2 T \cdot \mathbf{t}(x, t) \quad (71)$$

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma_{xt}}{\partial t^2} = \mathbf{T}(x + \delta x, t) - \mathbf{T}(x, t) + \mathbf{G} \cdot \rho \cdot \delta x, \quad (72)$$

holds for a every point $(x, t) \in (a, b) \times [0, +\infty)$ and for every $\delta x \in D$. Finally, let us suppose that

$$\frac{\partial \varphi}{\partial x}(x, t) \text{ is invertible.}$$

Then at this point (x, t) the following sentences are equivalent

1. $\rho(x, t) \cdot \frac{\partial^2 \gamma_2}{\partial t^2}(x, t) =_2 T \cdot \frac{\partial^2 \gamma_2}{\partial x^2}(x, t) + G_2(x, t) \cdot \rho(x, t)$
2. $\varphi(x, t) \in I_4$.

Finally, if (2) holds for every $(x, t) \in (a, b) \times [0, +\infty)$, then

$$\text{length}(\gamma_{-,t}) =_2 b - a.$$

To simplify the proof of this result, we need two lemmas.

Lemma 42. Let $a, b \in \mathbb{R}$ with $a < b$ and let $f, g : (a, b) \longrightarrow \bullet\mathbb{R}$ be non standard smooth functions such that

$$f(x) =_2 g(x) \quad \forall x \in (a, b).$$

Then

$$f(x + h) - f(x) = g(x + h) - g(x) \quad \forall h \in D \quad \forall x \in (a, b)$$

Lemma 43. Let $m, h \in \bullet\mathbb{R}$, and suppose that m is invertible and $0 \leq h \leq \pi$, then the following properties are equivalent:

$$1. \quad m \cdot \cos^3 h =_2 m$$

$$2. \quad h \in I_4.$$

Proof of Theorem 41: We firstly note that, assuming (70), the tangent vector $\mathbf{t}(x, t)$ always exists in $\bullet\mathbb{R}$. In fact we have $\frac{\partial\gamma_1}{\partial x}(x, t) = 1 + \alpha(t)$ so that both $\frac{\partial\gamma_1}{\partial x}(x, t)$ and $\left[\frac{\partial\gamma_1}{\partial x}(x, t)\right]^2 + \left[\frac{\partial\gamma_2}{\partial x}(x, t)\right]^2$ are invertible; we can hence take its square root and then the inverse to define the unit tangent vector. Now we prove that (1) implies (2). Let us take a generic $\delta x \in D$. Projecting (72) on \vec{e}_2 we get

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma_2}{\partial t^2} = \mathbf{T}(x + \delta x, t) \cdot \vec{e}_2 - \mathbf{T}(x, t) \cdot \vec{e}_2 + G_2 \cdot \rho \cdot \delta x.$$

But from (71) and because smooth operations preserve $=_2$, we get $\mathbf{T} \cdot \vec{e}_2 =_2 T \cdot \mathbf{t} \cdot \vec{e}_2$. Therefore, from Lemma 42 we obtain

$$\begin{aligned} \mathbf{T}(x + \delta x, t) \cdot \vec{e}_2 - \mathbf{T}(x, t) \cdot \vec{e}_2 &= T \cdot \mathbf{t}(x + \delta x, t) \cdot \vec{e}_2 - T \cdot \mathbf{t}(x, t) \cdot \vec{e}_2 \\ &= T \cdot \sin \varphi(x + \delta x, t) - T \cdot \sin \varphi(x, t) \\ \rho \cdot \delta x \cdot \frac{\partial^2 \gamma_2}{\partial t^2} &= T \cdot \sin \varphi(x + \delta x, t) - T \cdot \sin \varphi(x, t) + G_2 \cdot \rho \cdot \delta x. \end{aligned} \quad (73)$$

On the other hand, we can multiply (1) by δx (so that $=_2$ becomes $=$, see Theorem 40) obtaining

$$\begin{aligned} \rho \cdot \delta x \cdot \frac{\partial^2 \gamma_2}{\partial t^2} &= T \cdot \left[\frac{\partial \gamma_2}{\partial x}(x + \delta x, t) - \frac{\partial \gamma_2}{\partial x}(x, t) \right] + G_2 \cdot \rho \cdot \delta x \\ &= T \cdot \tan \varphi(x + \delta x, t) \cdot \frac{\partial \gamma_1}{\partial x}(x + \delta x, t) - T \tan \varphi(x, t) \cdot \frac{\partial \gamma_1}{\partial x}(x, t) + G_2 \cdot \rho \cdot \delta x, \end{aligned} \quad (74)$$

Equating (73) and (74) and canceling T we get

$$\sin \varphi(x + \delta x, t) - \sin \varphi(x, t) = \tan \varphi(x + \delta x, t) \cdot \frac{\partial \gamma_1}{\partial x}(x + \delta x, t) - \tan \varphi(x, t) \cdot \frac{\partial \gamma_1}{\partial x}(x, t)$$

$$\begin{aligned} \delta x \cdot \cos \varphi \cdot \frac{\partial \varphi}{\partial x} &= \delta x \cdot \frac{1}{\cos^2 \varphi} \cdot \frac{\partial \varphi}{\partial x} \cdot \frac{\partial \gamma_1}{\partial x}(x, t) + \tan \varphi \cdot \frac{\partial^2 \gamma_1}{\partial x^2}(x, t) \\ &= \delta x \cdot \frac{1}{\cos^2 \varphi} \cdot \frac{\partial \varphi}{\partial x} \cdot [1 + \alpha(t)] \\ &= \delta x \cdot \frac{1}{\cos^2 \varphi} \cdot \frac{\partial \varphi}{\partial x} \end{aligned} \quad (75)$$

where, as usual, every function, if not otherwise indicated, is calculated at (x, t) . Let us note that, in (75) we have used the property $\delta x \cdot \alpha(t) = 0$ because $\delta x \in D$ and $\alpha(t) \in I_2$; moreover, from (51) if $\varphi = \frac{\pi}{2}$ we would have $\frac{\partial \gamma_2}{\partial x} \cdot \cos \varphi = 0 = \frac{\partial \gamma_1}{\partial x} \cdot \sin \varphi = 1 + \alpha(t)$, which is impossible because $\alpha(t) \in D_\infty$. Setting, for simplicity, $m := \frac{\partial \varphi}{\partial x}(x, t) \in \bullet \mathbb{R}$, from (75) and canceling δx , we have

$$m \cdot \cos^3 \varphi =_2 m, \quad (76)$$

By Lemma 43 this implies the conclusion.

Vice versa, if φ is an infinitesimal of order less than or equal 4, then by Lemma 43 we obtain (76) and we can go over again the previous passages in the opposite direction to prove (1).

Now, let us suppose that $\varphi(x, t) \in I_4$ for every $(x, t) \in (a, b) \times [0, +\infty)$, then

$$\begin{aligned} \text{lenght}(\gamma_{-,t}) &= \int_a^b \sqrt{[1 + \alpha(t)]^2 + \left[\frac{\partial \gamma_2}{\partial x}(x, t) \right]^2} dx \\ &= \int_a^b \sqrt{1 + 2\alpha(t) + \left[\frac{\partial \gamma_2}{\partial x}(x, t) \right]^2} dx, \end{aligned} \quad (77)$$

because $\alpha(t) \in I_2$ and hence $\alpha(t)^2 = 0$. But $[1 + \alpha(t)] \cdot \sin \varphi = \frac{\partial \gamma_2}{\partial x}(x, t) \cdot \cos \varphi$, so

$$\begin{aligned} \frac{\partial \gamma_2}{\partial x}(x, t) &= [1 + \alpha(t)] \tan \varphi \\ &= [1 + \alpha(t)] \left(\varphi + \frac{\varphi^3}{3} \right) \\ &= \varphi + \frac{\varphi^3}{3} + \alpha(t) \cdot \varphi, \end{aligned}$$

because $\alpha(t) \in I_2$ and $\varphi \in I_4$ and hence $\alpha(t) \cdot \varphi^3 = 0$. Substituting this in (77) and using the derivation formula for the function $x \mapsto \sqrt{1+x}$ we obtain

$$\begin{aligned} \sqrt{1 + 2\alpha(t) + \left[\frac{\partial \gamma_2}{\partial x}(x, t) \right]^2} &= 1 + \frac{1}{2} \cdot \left\{ 2\alpha(t) + \left[\frac{\partial \gamma_2}{\partial x}(x, t) \right]^2 \right\} \\ &= 1 + \alpha(t) + \frac{1}{2} \left[\varphi + \frac{\varphi^3}{3} + \alpha(t) \cdot \varphi \right]^2 \\ &= 1 + \alpha(t) + \frac{\varphi^2}{2} + \frac{\varphi^4}{3} + \alpha(t) \cdot \varphi^2. \end{aligned}$$

Therefore

$$\begin{aligned} \text{lenght}(\gamma_{-,t}) &= \int_a^b \left[1 + \alpha(t) + \frac{\varphi(x, t)^2}{2} + \frac{\varphi(x, t)^4}{3} + \alpha(t) \cdot \varphi(x, t)^2 \right] dx \\ &= b - a + \alpha(t) \cdot (b - a) + \int_a^b \left[\frac{\varphi(x, t)^2}{2} + \frac{\varphi(x, t)^4}{3} + \alpha(t) \cdot \varphi(x, t)^2 \right] dx. \end{aligned} \quad (78)$$

Using the Theorem 38 it is not hard to prove that the last integral in (78) is an infinitesimal of order less than or equal 2, so the conclusion follows from the hypothesis $\alpha(t) \in I_2$. ■

Proof of Lemma 42: First of all, from the hypothesis $f(x) =_2 g(x)$ for every $x \in (a, b)$, we get that

$$^\circ f(x) = ^\circ g(x) \quad \forall x \in (a, b). \quad (79)$$

Now, let us fix a point $x \in (a, b)$. From Theorem 38 we obtain that we can write

$$\begin{aligned} f(x_1) &= a_0(x_1) + \sum_i p_i \cdot a_i(x_1) \\ g(x_1) &= b_0(x_1) + \sum_j q_j \cdot b_j(x_1), \end{aligned}$$

for every $x_1 \in (x - \delta, x + \delta) \subseteq (a, b)$ and where $p_i, q_j \in D_\infty$ and a_i, b_j are ordinary smooth functions defined in an open neighbourhood V of $^\circ x \in (a, b) \cap \mathbb{R}$. From (79) we have $a_0(^\circ x_1) = b_0(^\circ x_1)$ for every $x_1 \in \bullet V$ so that $a_0 = b_0$ on V and hence also $\bullet a_0 = \bullet b_0$ on $\bullet V$. Therefore

$$f(r) - g(r) = \sum_i p_i \cdot a_i(r) - \sum_j q_j \cdot b_j(r) \quad \forall r \in (a, b) \cap \mathbb{R}. \quad (80)$$

This difference must have order less than or equal 2 because $f(r) =_2 g(r)$, so

$$\omega \left[\sum_i p_i \cdot a_i(r) - \sum_j q_j \cdot b_j(r) \right] = \max_i \omega [p_i \cdot a_i(r)] \vee \max_j \omega [q_j \cdot b_j(r)] \leq 2.$$

Let us suppose, for simplicity, that $\omega(p_1 \cdot a_1(r))$ is this term of maximum order. Because $a_1(r) \in \mathbb{R}$ it must be that $\omega(p_1) \leq 2$ and hence also $\omega(p_i) \leq \omega(p_1) \leq 2$ and $\omega(q_j) \leq \omega(p_1) \leq 2$. Finally we have

$$\begin{aligned} f(x+h) - f(x) &= h \cdot f'(x) \\ &= h \cdot a'_0(x) + \sum_i h \cdot p_i \cdot a'_i(x), \end{aligned}$$

but $a'_0(x) = b'_0(x)$ because $a_0 = b_0$ and $h \cdot p_i = 0$ because $\omega(h) < 2$ and $\omega(p_i) \leq 2$; we hence obtain

$$\begin{aligned} f(x+h) - f(x) &= h \cdot b'_0(x) \\ &= h \cdot b'_0(x) + \sum_j h \cdot q_j \cdot b'_j(x) \\ &= h \cdot g'(x) \\ &= g(x+h) - g(x). \end{aligned}$$

■

Proof of Lemma 43: If $m \cdot \cos^3 h =_2 m$, then the standard parts of both sides must be equal

$$^\circ(m \cdot \cos^3 h) = ^\circ m$$

$$^\circ m \cdot \cos^3(^\circ \varphi) = ^\circ m.$$

By hypotheses m is invertible, hence $^\circ m \neq 0$ and we obtain that $^\circ h = 0$ because $0 \leq h \leq \pi$, i.e. $h \in D_\infty$. Moreover, from infinitesimal Taylor's formula applied to $\cos h$, and from $m \cdot \cos^3 h =_2 m$ we obtain

$$\begin{aligned} m \cdot \left(1 - \sum_{1 \leq i < \frac{\omega(h)+1}{2}} (-1)^i \frac{h^{2i}}{(2i)!} \right)^3 &= _2 m \\ m \cdot (1 + a \cdot h^2)^3 &= _2 m \\ m \cdot (1 + a^3 h^6 + 3a h^2 + 3a^2 h^2) &= _2 m \\ m \cdot (1 + \alpha \cdot h^2) &= _2 m \end{aligned}$$

where $a := -\sum_{1 \leq i < \frac{\omega(h)+1}{2}} (-1)^i \frac{h^{2i-2}}{(2i)!} \in {}^\bullet \mathbb{R}$ and $\alpha := 3a^2 + 3a + a^3 h^4$ are invertible Fermat reals. From this we get $m \cdot \alpha \cdot h^2 =_2 0$ and hence $h^2 =_2 0$, i.e. $\omega(h^2) \leq 2$ and hence $\omega(h) \leq 4$.

Vice versa, if h is an infinitesimal of order less than or equal 4 (so that $\varphi^n = 0$ if $n \geq 5$) we have

$$\begin{aligned} \cos^3 h &= \left(1 - \frac{h^2}{2} + \frac{h^4}{4!} \right)^3 = \\ &= 1 - 3\frac{h^2}{2} + 3\frac{h^4}{4!}. \end{aligned}$$

Therefore, $m \cdot \cos^3 h = m - 3mh^2 \cdot \left(\frac{1}{2} - 3\frac{h^2}{4!} \right)$ so that $m \cdot \cos^3 h - m = -3mh^2 \cdot \left(\frac{1}{2} - 3\frac{h^2}{4!} \right)$ is an infinitesimal of order $\omega(h^2) \leq 2$, i.e. $m \cos^3 h =_2 m$.

■

The reader with a certain knowledge of SDG had surely noted that this deduction of the wave equation cannot be reproduced in SDG because of the use of non standard smooth functions, of the use of equalities up to k -th order infinitesimals and because of the frequent use of the useful Theorem 12 to decide products of powers of nilpotent infinitesimals.

XV. CONCLUSIONS

The problem to turn informal infinitesimal methods into a rigorous theory has been faced by several authors. The most used theories, i.e. NSA and SDG, require a good knowledge of Mathematical Logic and a strong formal control. Some others, like Weil functors (see e.g. Kriegl and Michor [21]) or the Levi-Civita field (see e.g. Shamseddine [25]) are mainly based on formal/algebraic methods and sometimes lack the intuitive meaning. In this initial work, we have shown that it is possible to bypass the inconsistency of SIA with classical logic modifying the Kock-Lawvere axiom (see e.g. Lavendhomme [22]) and keeping always a very good intuitive meaning. We have seen how to define the algebraic operations between this type of nilpotent infinitesimals, infinitesimal Taylor formula and order properties. In the final part we have seen several elementary examples of the use of these infinitesimals, some of them taken from classical deductions of elementary Physics. In our opinion, these examples are able to show that some results that frequently may appear as unnatural in a standard context, using Fermat reals can be discovered, even by suitably designed algorithm. Moreover, our generalization of the classical proof of the wave equation have shown that a rigorous theory of infinitesimals permits to obtain results that are not accessible using only an intuitive approach.

-
- [1] S. Albeverio, J.E. Fenstad, R. Høegh-Krohn, and T. Lindstrøm. *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*. Pure and Applied Mathematics. Academic Press, 1988. 2nd ed., Dover, 2009.
 - [2] E.T. Bell. *Men of Mathematics*. Simon and Schuster, New York, 1937.
 - [3] J.L. Bell. *A Primer of Infinitesimal Analysis*. Cambridge University Press, 1998.
 - [4] V. Benci and M. Di Nasso. A ring homomorphism is enough to get nonstandard analysis. *Bull. Belg. Math. Soc. - S. Stevin*, 10:481–490, 2003.
 - [5] V. Benci and M. Di Nasso. A purely algebraic characterization of the hyperreal numbers. *Proceedings of the American Mathematical Society*, 133(9):2501–05, 2005.
 - [6] W. Bertram. *Differential Geometry, Lie Groups and Symmetric Spaces over General Base Fields and Rings*. American Mathematical Society, Providence, 2008.
 - [7] M. Berz. *Analysis on a Nonarchimedean Extension of the Real Numbers*. Mathematics Summer Graduate School of the German National Merit Foundation, MSUCL-933, Department of Physics, Michigan State University, 1992 and 1995 edition, 1994.
 - [8] T. Bröcker. *Differentiable germs and catastrophes*, volume 17 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1975.
 - [9] J.H. Conway. *On Numbers and Games*. Number 6 in L.M.S. monographs. Academic Press, London & New York, 1976.
 - [10] P.A.M. Dirac. *General Theory of Relativity*. John Wiley and Sons, 1975.
 - [11] C.H. Edwards. *The Historical Development of the Calculus*. Springer-Verlag, New York, 1979.
 - [12] C. Ehresmann. Les prolongements d’une variété différentiable: Calculus des jets, prolongement principal. *C. R. Acad. Sc. Paris*, 233:598–600, 1951.
 - [13] A. Einstein. *Investigations on the Theory of the Brownian Movement*. Dover, 1926.
 - [14] H. Eves. *An Introduction to the History of Mathematics*. Saunders College Publishing, Fort Worth, TX, 1990.
 - [15] P. Giordano. Infinitesimal differential geometry. *Acta Mathematica Universitatis Comenianae*, LXIII(2):235–278, 2004.
 - [16] P. Giordano. Fermat reals: Nilpotent infinitesimals and infinite dimensional spaces. arXiv:0907.1872, July 2009.

- [17] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities*, volume 14 of *Graduate texts in mathematics*. Springer, Berlin, 1973.
- [18] A. Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, volume 19 of *Frontiers in Applied Mathematics*. SIAM, 2000.
- [19] C.W. Henson. Foundations of nonstandard analysis. a gentle introduction to nonstandard extension. In L.O. Arkeryd, N.J. Cutland, and C.W. Henson, editors, *Nonstandard analysis: theory and applications (Edinburgh, 1996)*, pages 1–49, Dordrecht, 1997. NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., vol. 493, Kluwer Acad. Publ.
- [20] A. Kock. *Synthetic Differential Geometry*, volume 51 of *London Math. Soc. Lect. Note Series*. Cambridge Univ. Press, 1981.
- [21] A. Kriegl and P.W. Michor. Product preserving functors of infinite dimensional manifolds. *Archivum Mathematicum (Brno)*, 32, 4:289–306, 1996.
- [22] R. Lavendhomme. *Basic Concepts of Synthetic Differential Geometry*. Kluwer Academic Publishers, Dordrecht, 1996.
- [23] I. Moerdijk and G.E. Reyes. *Models for Smooth Infinitesimal Analysis*. Springer, Berlin, 1991.
- [24] G. Prodi. *Analisi matematica*. Ed. Bollati Boringhieri, Torino, 1970.
- [25] K. Shamseddine. *New Elements of Analysis on the Levi-Civita Field*. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1999.
- [26] G.E. Silov. *Analisi matematica. Funzioni di una variabile* (it. transl.). Mir, Mosca, 1978.
- [27] C. Truesdell. *A First Course in Rational Continuum Mechanics: v.1 General Concepts, 2nd ed.*, volume 71 of *Pure and Applied Mathematics*. Academic Press Inc., 1991.
- [28] V.S. Vladimirov. *Equazioni della fisica matematica*. MIR, 1987.